TIINA RIIHIMÄKI

Modification of Structural and Functional Properties of Avidin Protein Using Targeted Random Mutagenesis

■

ACADEMIC DISSERTATION
To be presented, with the permission of
the board of the Institute of Biomedical Technology of the University of Tampere,
for public discussion in the Auditorium of Finn-Medi 5,
Biokatu 12, Tampere, on September 23rd, 2011, at 12 o'clock.

UNIVERSITY OF TAMPERE

UNIVERSITY
OF TAMPERE

ACADEMIC DISSERTATION
University of Tampere, Institute of Biomedical Technology
VTT Technical Research Center of Finland, Molecular Diagnostic and Immunotechnology
Tampere Graduate Program in Biomedicine and Biotechnology (TGPBB)
Finland

Supervised by
Professor Markku S. Kulomaa
University of Tampere
Finland
Docent Vesa P. Hytönen
University of Tampere
Finland

Reviewed by
Professor Kari Airenne
University of Eastern Finland
Finland
Doctor of Philosophy Petri Saviranta
University of Turku
Finland

Distribution
Bookshop TAJU
P.O. Box 617
33014 University of Tampere
Finland

Tel. +358 40 190 9800
Fax +358 3 3551 7685
taju@uta.fi
www.uta.fi/taju
http://granum.uta.fi

Cover design by
Mikko Reinikka

Acta Universitatis Tamperensis 1641
ISBN 978-951-44-8531-2 (print)
ISSN-L 1455-1616
ISSN 1455-1616

Acta Electronica Universitatis Tamperensis 1103
ISBN 978-951-44-8532-9 (pdf)
ISSN 1456-954X
http://acta.uta.fi

Tampereen Yliopistopaino Oy – Juvenes Print
Tampere 2011

CONTENTS

LIST OF ORIGINAL COMMUNICATIONS

This thesis is based on the following original communications that are referred to the text by their Roman numerals (I-IV). All papers are reproduced with permission of the respective copyright holders.

I. Repo S, **Paldanius TA**, Hytönen VP, Nyholm TK, Halling KK, Huuskonen J, Pentikäinen OT, Rissanen K, Slotte JP, Airenne TT, Salminen TA, Kulomaa MS, Johnson MS (2006): Binding properties of HABA-type azo derivatives to avidin and avidin-related protein 4. Chem Biol. 2006 Oct;13(10):1029-39.

II. **Riihimäki TA**, Kukkurainen S, Varjonen S, Hörhä J, Nyholm TKM, Kulomaa MS, Hytönen VP (2011): Construction of chimeric dual-chain avidin by tandem fusion of the related avidins. Plos One 2011 May; 6(5):e20535-e20535. Epub 2011 May 31.

III. Heikkinen JJ, **Riihimäki TA**, Määttä JAE, Suomela SE, Kantomaa J, Kulomaa MS, Hytönen VP, Hormi OEO: Covalent biofunctionalization of cellulose acetate with thermostable chimeric avidin. ACS Appl Mater Interfaces. 2011 Jul 27; 3(7):2240-2245. Epub 2011 Jun 16.

IV. **Riihimäki TA**, Hiltunen S, Rangl M, Nordlund HR, Määttä JAE, Ebner A, Hinterdorfer P, Kulomaa MS, Takkinen K, Hytönen VP: Modification of the loops in the ligand-binding site turns avidin into a steroid-binding protein. BMC Biotechnol. 2011 Jun 9;11(1):64.

ABBREVIATIONS

3D	three-dimensional
AVR	avidin-related gene
AVR	avidin-related protein
BSA	bovine serum albumin
cDNA	complementary DNA
CDR	complementarity determing region
C_H	constant region, heavy chain
C_L	constant region, light chain
c-Myc	proto-oncogene
CTLA-4	cytotoxic T-lymphocyte antigen-4
Da	Dalton
DIRE	directed interaction rescue
DSC	differential scanning calorimetry
EDTA	ethylenediaminetetraacetic acid
ELISA	enzyme-linked immunosorbent assay
FDA	food and drug administration
FLAG	protein tag
HABA	2-(4'-hydroxyazobenzene) benzoic acid
HAMA	human anti-mouse antibody
HARA	human anti-rat antibody
HER2	human epidermal growth factor receptor 2
HIV-1	human immunodeficiency virus 1
Ig	immunoglobulin
K_D	dissociation constant
K_{ass}	association rate constant
K_{diss}	dissociation rate constant
LB	lysogeny broth
M	molar concentration

p3	phage coat protein 3
p8	phage coat protein 8
PCR	polymerase chain reaction
PDB	protein data bank
PEG	polyethylene glycol
SAP	selection and amplification of phage
SDS-PAGE	sodium-dodecyl polyacrylamide gel electrophoresis
SIP	selectively-infective phage
T_m	transition midpoint of heat denaturation
TNF	tumor necrosis factor
VEGF	vascular endothelial growth factor
V_H	variable region, heavy chain
V_L	variable region, light chain
wt	wild type
Å	ångström

YHTEENVETO

Uusien proteiinien kehittäminen biokemiallisiin ja lääketieteellisiin sovellutuksiin on kehittynyt nopeasti viime vuosina. Nykyteknologia mahdollistaa esimerkiksi sellaisten proteiinirakenteiden valmistamisen, joita ei esiinny luonnossa, sekä proteiinien laaja-alaisen muokkaamiseen bioteknologisiin sovellutuksiin paremmin sopiviksi. Moderneilla tekniikoilla voidaan parantaa proteiinin affiniteettia ja spesifisyyttä sekä jopa lisätä proteiineille uusia toimintoja. Lisäksi proteiinien tuottotasoja voidaan kasvattaa esimerkiksi lisäämällä proteiinien kestävyyttä tai liukoisuutta. Tutkimusmenetelmien kehittyessä ominaisuuksiltaan muokattujen proteiinien tarve kasvaa kokoajan. Erityisen mielenkiinnonkohteena ovat molekyylit, jotka kykenevät tarkkaan ja spesifiin molekyylitunnistukseen. Tunnistusmolekyylien valmistamisen lisäksi proteiinien muokkaus tarjoaa arvokasta tietoa proteiinien vuorovaikutuksista sekä proteiinin rakenteen ja toiminnan yhteydestä.

Tämän väitöskirjatutkimuksen tavoitteena oli kehittää uusia muokattuja avidiinimolekyylejä avidiini-biotiiniteknologiaan ja samalla saada lisää tietoa avidiinin ligandin sitomisominaisuuksista. Tutkimuksessa avidiinin kolmiulotteinen rakenne määritettiin korkealla resoluutiolla, jolloin havaittiin avidiinin tynnyrirakenteen beta-säikeiden 3 ja 4 välissä olevan silmukan tärkeys ligandin sitomisessa. Analyyseissä, joissa määritettiin avidiinin kykyä sitoutua atsoväreihin, avidiini ja avidiinin kaltainen proteiini 4 (AVR4) osoittivat toisistaan poikkeavaa ligandin sitomiskykyä. Eroja tutkittiin molekyylimallinnuksen avulla käyttäen ns. molekyylien telakoitumista. Mallinnus osoitti kolmiulotteisen rakenteen tavoin beta-säikeiden 3 ja 4 välissä olevan silmukan tärkeyden ligandin sitomisessa.

Väitöskirjatutkimuksessa avidiinin rakennetta muokattiin laajalti, jotta saataisiin sopivampia molekyylejä diagnostisiin ja biokemiallisiin sovellutuksiin. Alunperin kehitetyn kaksoisketjuavidiinin käytettävyyttä parannettiin kehittämällä kimeerinen kaksoisketjuavidiini avidiinin sirkulaarisesta permutantista ja AVR4:n sirkulaarisesta permutantista. Kimeerisellä fuusioproteiinilla havaittiin parantuneet proteiinin tuottotasot sekä lämmönkestävyys verrattuna alkuperäiseen kaksoisketjuavidiiniin.

Lisäksi proteiinia koodaavan DNA-ketjun monistaminen PCR-menetelmällä oli tehokkaampaa kimeerisen fuusiogeenin avulla. Tämä mahdollistaa mm. satunnaismutageneesimenetelmän käytön molekyylin edelleen kehittämisessä.

Sirkulaaristen permutaatioiden lisäksi tutkimuksessa modifioitiin myös avidiiniproteiinin ligandin sitomiseen osallistuvaa silmukka-aluetta kohdennetun satunnaismutageneesin avulla. Tavoitteena oli tuottaa avidiini, jolla on heikentynyt sitomiskyky avidiinin luonnolliseen ligandiin biotiiniin ja toisaalta korkea sitomisaffiniteetti uuteen pienmolekyyliin. Valmistettujen avidiinigeenikirjastojen seulontaan käytettiin faagidisplay-menetelmää. Seulontamenetelmää varten avidiiniproteiinia ilmennettiin M13-bakteriofaagin pinnalla p3-pintaproteiinin fuusiona. Seulonnan tuloksena löydettiin lämmönkestäviä avidiinimutantteja, joilla on mikromolaarinen affiniteetti steroidihormoni testosteroniin. Lisäksi valikoidun avidiinimutantin affiniteettia biotiinin saatiin heikennettyä kohdennetulla mutageneesillä niin, että proteiini sitoi mieluumin steroideja kuin biotiinia.

Tutkimuksessa edistettiin avidiini-biotiini-teknologiaa kehittämällä myös bioaktiivisia kalvoja. Sovelluksessa yhdistettiin aikaisemmin kehitetyn lämmönkestävän kimeerisen avidiinin ominaisuudet selluloosa-asetaattikalvojen helppokäyttöisyyteen. Kalvot päällystettiin kimeerisellä avidiinilla ja kalvojen bioaktiivisuutta seurattiin tritium-biotiinilla. Kalvojen bioaktiivisuus säilyi lähes muuttumattomana kolmen kuukauden pituisen kokeen aikana, vaikka kalvot säilytettiin huoneenlämpötilassa. Kehitettyjä kalvoja voitaisiin käyttää perustana useissa sovelluksissa, esimerkiksi diagnostisten pikatestien kehityksessä.

ABSTRACT

In recent years, the development of novel scaffolds from non-immunoglobulin proteins have been an area of great intrest, and a powerful new technology has been born. Today, it is possible to design structures never seen before in nature and to generate proteins with improved suitability for biomedical and biochemical applications. With modern techniques, the affinity and specificity of target proteins can be enhanced, and novel functions or activities can even be added to modified proteins. Additionally, proteins with more stable structures or higher solubility have been created, thus increasing the expression of recombinant proteins. Ongoing research may answer the increasing need for molecules capable of more sensitive and specific recognition in biotechnology applications. Further research will provide valuable knowledge about molecular recognition and protein structure-function relationships.

The aim of this study was to develop novel molecules for avidin-based technologies and, at the same time, to acquire more information about the ligand-binding properties of avidin. A high-resolution 3D -structure of avidin was determined, and it revealed an important role for the loop between beta-strands 3 and 4 in ligand binding. Analyses of the binding of avidin and avidin-related protein 4 (AVR4) to azo dyes showed clearly different binding modes for each protein. Therefore, molecular docking was used to analyze the difference in binding. This experiment also highlighted the importance of the loop between beta-strands 3 and 4 in ligand-binding.

In this study, the avidin protein was extensively modified to be more suitable for diagnostic and biochemical applications. The utility of the original dual-chain avidin protein was improved by constructing chimeric dual-chain avidin proteins consisting of a fusion of circularly permutated wild-type (wt) avidin and circularly permuted AVR4. The chimeric tandem fusion showed enhanced protein expression and thermal stability compared to the original dual-chain avidin. Moreover, PCR amplifica-

tion was more straightforward for the chimeric protein. Interestingly, binding analyses of the chimeric tandem fusion showed heterogeneous biotin-binding.

In addition to circular permutation, the loop areas in the ligand-binding site of wt avidin were modified in this study. The loop area was subjected to targeted random mutagenesis, and genes carrying the desired mutations were selected from a library by the phage display method. For selection, avidin was displayed on the M13 phage as a fusion with the coat protein p3. As a result, thermostable avidin variants with micromolar affinity for the steroid hormone testosterone were obtained. In terms of ligand binding, the steroid-binding avidin preferred steroids to its natural ligand, biotin.

To further advance the possibile applications of avidin-biotin -technology, avidin-coated cellulose acetate-films were developed. Films covered with thermostable chimeric avidin and stored at room temperature remained biologically active during three-month test period. These films could be used as a universal base for various applications, including in diagnostic platforms.

INTRODUCTION

Proteins can be found everywhere, and they are essential for many biological processes. Inside the cell, proteins work together with other biological macromolecules, such as polysaccharides and nucleic acids. Proteins vary considerably in their structures and functions. Many proteins are enzymes that catalyze biochemical reactions, whereas proteins such as actin and myosin have mechanical functions. Proteins capable of molecular recognition are essential for cell signaling and for the immune response.

Generally, the best known class of recognition molecules is antibodies, also known as immunoglobulins (Igs). Antibodies are Y-shaped molecules that participate in the immune response by identifying and neutralizing foreign intruders. A molecule recognized by antibodies is referred to as an antigen. Traditionally, antibodies have been involved in many biomedical applications because of their characteristic ability to recognize and bind other molecules. Antibodies are a true success story; today more than 20 different antibodies have been approved for therapy in Europe and in the USA (Gebauer, Skerra 2009). There are many reasons for this success. One important reason is the ability to generate antibodies that target a wide range of molecules, often with picomolar affinities.

However, due to the complex structure of antibodies, the engineering of proteins outside the Ig family began two decades ago, and nowadays, the development of novel proteins has grown into a powerful technology. With developed technologies, it is possible to design structures never seen before in nature and catalyze reactions for which no natural enzyme exists (Smith, Hecht 2011). Additionally, the isolation of sequences that have no biological ancestors but nonetheless enable the growth of living cells is possible (Smith, Hecht 2011). These advances led to the production of proteins more suitable for biomedical and biochemical applications. The affinity and specificity of several proteins toward their ligands have been enhanced, and novel functions or activities have been generated. Proteins have also been tailored to produce more stable molecules or to improve solubility, which helps to increase pro-

duction levels. For therapeutic purposes, the half-lives of protein molecules have been extended, and the immunogenicity of many molecules has been decreased (Kurtzman *et al.* 2001; Leader, Baca & Golan 2008).

Generally, proteins that recognize other molecules with high specificity can be used for diagnostic applications *in vivo* and in medical applications, such as targeted therapeutic drug treatment. Recombinant proteins are highly specific and may provide effective treatment without the need for gene delivery. Protein-based therapeutics have less potential to interfere with normal biological processes and are therefore often well tolerated (Leader, Baca & Golan 2008). Besides therapeutic uses, proteins with tailored affinities can be included in several applications, such as bioseparation, detection and proteomic analysis (Gronwall, Stahl 2009). Currently, some of the most promising non-Ig scaffolds are Staphylococcal protein A, lipocalins, a fibronectin domain, an ankyrin consensus repeat domain, and thioredoxin (Skerra 2007).

At present, recombinant human proteins, including monoclonal antibodies, interferons, vaccines, hormones and modified enzymes make up the majority of FDA-approved biotechnology-based medicines (Leader, Baca & Golan 2008). Personalized approaches for the treatment of disease are becoming more common all the time. These approaches require more specific properties from the recognition molecules involved. Non-Ig scaffolds with specific recognition properties will provide considerable alternatives to novel approaches. In addition to generating novel molecules, protein engineering gives us valuable knowledge about molecular recognition and protein structure-function relationships.

This review of the literature will focus on the generation of proteins with novel affinities. One of the best examples of such proteins is antibodies due to their considerable influence and importance for the development of methods and applications in the field. Beyond antibodies, the most successful non-Ig-based recognition molecules are also discussed. Additionally, the strategies for gene library construction and methods for selection are described. In the experimental section, the protein engineering methods used to modify the chicken avidin protein to be more suitable for life-science applications are detailed.

2. REVIEW OF THE LITERATURE

2.1 Antibodies are naturally evolved affinity proteins

Antibodies, also known as immunoglobulins (Igs), are a central part of the humoral immune system. The immune system protects the body against foreign intruders, such as bacteria, viruses, fungi and parasites. Upon activation of immune system, antibodies are produced by the B lymphocytes. Antibodies are also known to participate in allergic reactions and autoimmune diseases.

Unmodified polyclonal antibodies were the first molecules accepted for therapeutic use, and today, antibodies are the most successful biomolecules utilized in biotechnological and medical applications. Mouse hybridomas (Kohler, Milstein 1975) were the first reliable source of monoclonal antibodies developed for a number of therapeutic applications. B cell hybridoma technology offers a reproducible supply of an antibody with a high specificity (Nelson *et al.* 2000), but unfortunately, if treatment is repeated, mouse-derived molecules provoke an anti-mouse response in humans. Therefore, different strategies have been developed to redirect, mask or avoid human immune surveillance. For instance, chimeric antibodies in which the variable regions of the mouse antibody are fused to the constant regions of the human antibody have been generated (Boulianne, Hozumi & Shulman 1984; Morrison *et al*. 1984). In humanized antibodies, only the complementarity-determining regions (CDRs) of the variable regions are of non-human origin. Throughout the years, technologies have evolved, and fully synthetic human antibodies having high affinities for antigens such as DNA, peptides, and proteins have also been generated (Knappik *et al.* 2000, Lonberg 2008). Most recently, innovative structural designs have improved *in vivo* pharmacokinetics, expanded immune repertoires and permitted screening using refractory targets and complex proteome arrays. These *in vitro* evolution strategies have improved the affinity, specificity and expression levels of antibodies (Hudson, Souriau 2003). Engineered antibodies that are easily and inex-

pensively produced are widely used because of their capacity to detect multiple epitopes (Nelson *et al.* 2000).

2.1.1 Antibody structure

Generally speaking, antibodies are Y-shaped glycoprotein molecules that belong to the immunoglobulin superfamily (Schroeder, Cavacini 2010). However, antibodies can differ in structure; for example, the extent of glycosylation varies. Antibodies can be divided into five isotypes (Table 1): IgA, IgD, IgE, IgG, and IgM.

Table 1. Properties of immunoglobulin isotypes and subclasses. The table adapted adapted and modified from Schroeder and Cavacini (2010). Only the IgG isotype can cross the placenta and reach the fetus.

Serum (%)		Structure	Placental crossing	Other functions
IgG	75	Monomer	+	For all IgGs:
IgG1		Monomer	+	Secondary response
IgG2		Monomer	+	Neutralization of toxins and
IgG3		Monomer	+	viruses
IgG4		Monomer	+	
IgM	10	Pentamer	-	Primary response
IgA	15	Monomer, dimer	-	Mucosal response
IgA1		Monomer, dimer	-	
IgA2		Monomer, dimer	-	
IgD	<0.5	Monomer	-	Homeostasis
IgE	<0.01	Monomer	-	Allergy

IgG is the central antibody that recognizes and neutralizes pathogens. IgG also crosses the placenta to provide passive immunity to the fetus. IgGs can be split to four subclasses: IgG1, IgG2, IgG3, and IgG4 (Schroeder, Cavacini 2010). Like IgG, IgM also serves to help eliminate pathogens, but IgM also has the ability to facilitate the removal of apoptotic cells (Ehrenstein, Notley 2010). IgA can be found in mucosal areas, such as the gut, respiratory tract, and urogenital tract (Underdown, Schiff 1986). IgAs can be split to two subclasses: IgA1 and IgA2 (Schroeder, Cavacini 2010). IgD fine-tunes humoral responses and modulates B cell selection and homeostasis (Geisberger, Lamers & Achatz 2006). IgE acts during allergenic reaction and also protects against parasitic worms.

Antibodies are built from two identical heavy chains (50 to 70 kDa) and two identical light chains (23 kDa each) (Figure 1). The heavy and light chains are held together by disulphide bonds and by non-covalent interactions, but intra-chain disulfide bonds can also be found from the antibody structure. There are two gene families, κ and λ, for light chains, and each light chain consists of one variable (V_L) domain and one constant (C_L) domain. The heavy chain includes one variable (V_H) domain and three constant (C_H) domains. Each variable or constant domain is, on average, 12,000 to 13,000 Da in size and consists of 110 to 130 amino acids (Schroeder, Cavacini 2010).

In the center of the antibody molecule, there is a the hinge region that provides flexibility to the molecule. Upon protease treatment, the antibody molecule can be digested into smaller parts. The Fab fragment contains one V_L domain and one V_H domain with one C_L and one C_H domain. The Fc fragment of the antibody includes four C_H domains and specifies effector functions. The glycans associated with the Fc domain are known to affect the function of antibody (Schroeder, Cavacini 2010).

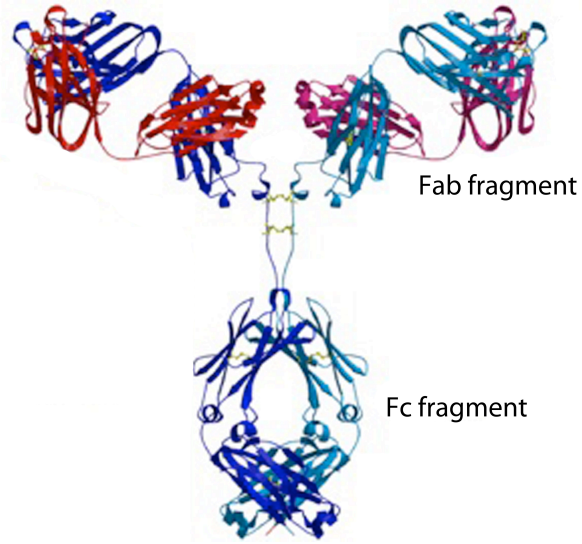


Figure 1. The structure of an IgG molecule. In the figure light chains are presented with red color and heavy chains in blue color. Disulphide-bridges are displayed as a yellow line in the figure. Figure from: http://www.biochem.arizona.edu/classes/bioc471/pages/Lecture10/Lecture10.html (9.6.2011).

2.1.1.1 The antigen recognition site of an antibody

The structural mechanism by which antibodies recognize their antigens is well studied and rather understood. Molecular recognition and ligand-binding occurs in the CDR, which is located at the very top of the antibody molecule. Six CDRs (CDR-V1 to V3 and CDR-H1 to H3) are located in the variable regions (Chothia *et al.* 1989). These loops are highly diverse in terms of amino acid composition and are in direct contact with the antigen. However, the key determinants of specificity in antigen binding are the CDR 3 loops, especially V_H, which contacts the antigen tightly (Xu, Davis 2000).

The overall composition of a functional CDR is biased in favor of certain amino acid residues. Sidhu *et al.* studied antigen binding by generating randomized CDRs with codons encoding only four amino acids (tyrosine, alanine, aspartate, and serine) (Fellouse, Wiesmann & Sidhu 2004). They discovered that antigen recognition was mediated primarily by tyrosine residues, whereas alanine and serine residues have no contact with the antigen. Alanine and serine residues induced conformational flexibility of the binding site. Sidhu *et al.* also postulated that the overabundance of tyrosine in natural antigen-binding sites is a consequence of its large chemically active side chain, which is particularly well -suited for making contacts with an antigen (Fellouse, Wiesmann & Sidhu 2004; Fellouse *et al.* 2005). Surprisingly, highly specific antibodies could be built using an interface consisting only of two amino acids, tyrosine and serine (Fellouse *et al.* 2005). This result is consistent with the discovery that the most specific antibodies are rich in tyrosine in CDR-H3 part of the antibody (Birtalan *et al.* 2008). However, the duration of co-evolution affects the protein moieties that contribute to ligand binding by an antibody. For instance, tyrosine, tryptophan and arginine mediate ligand binding at co-evolved interfaces, whereas small amino acids such as glycine, serine, alanine and threonine are favored in the naïve loop of antibodies (Zemlin *et al.* 2003).

Generally, tryptophan and arginine amino acid residues are scarce in the naïve molecular recognition sites of antibodies. However, Sidhu and co-workers noticed that the arginine residue becomes more abundant when an antibody has undergone affinity maturation, and glycine residues have favorable effects on antigen recognition in high-affinity antibodies when present in the CDR-H3 domain (Birtalan *et al.* 2008). Conformational flexibility, not just the specific interaction to an antigen, is of

great importance during antigen recognition by the CDR loop. The amino acid residues of glycine, tyrosine and serine have been shown to be frequently present in antibody hot spots and are also frequently mutated during somatic hypermutation (Clark *et al.* 2006).

2.1.1.2 Vast diversity generated by somatic hypermutation

The immune system is capable of triggering efficient responses against almost every foreign intruder. More than 10^9 different B lymphocyte clones circulate through the body, providing an enormous diversity of antibodies. The immediate response toward an intruder usually involves low affinity, but the affinity for the antigen increases as the response progresses (Neuberger, Milstein 1995).

In humans the genetic mechanism behind the generation of high-affinity binding of numerous antigens is the random recombination of an inherited set of gene segments followed by hypermutation. During hypermutation, genes encoding variable regions are modified, and these genes are translated into thousands of different antibodies. Somatic hypermutation creates point -mutations in IgG V(D)J gene segments (Krebs, Goldstein & Kilpatrick 2011). These mutations are targeted to certain positions, called 'hotspots,' where amino acid modifications can efficiently lead to affinity maturation (Neuberger, Milstein 1995). However, there are also other possibilities for acquiring genetic diversity such as gene conversion, which is frequently utilized in the immune response of chickens (Reynaud *et al.* 1987). The molecular mechanisms behind the antibody diversity generation have also been adapted for the creation of binding proteins *in vitro* (Skerra 2003), as will be discussed later.

2.1.2 Engineered antibodies

Antibodies have been extensively engineered to reduce the size of the molecule or to create multivalent binding proteins. For many applications, only the functional antigen-binding region of an antibody is needed. Compared to intact antibodies, smaller fragments have beneficial characteristics, such as better pharmacokinetics and tissue penetration.

Small antibody molecules can be produced by proteolytic cleavage. Depending on the enzyme, cleavage will result in a single Fab_2 fragment or two Fab fragments (Figure 2). Similarly, Fv fragments may also be acquired, but microbial expression of single-chain Fv molecules (scFv, Figure 2) is currently favored (Hudson, Souriau 2003). However, Fab and scFv molecules are monovalent and often have fast off rates and poor retention times for their targets (Hudson, Souriau 2003). To overcome these limitations, small antibody molecules have been engineered into dimeric, trimeric or tetrameric conjugates (Figure 2). Multivalent recombinant antibodies provide high binding avidity with optimal size (60-120 kDa) and are capable of rapid tumor penetration without rapid renal clearance (Todorovska *et al.* 2001). In many applications, recombinant antibodies have been fused with various substances, such as radionuclides, toxins, enzymes and viruses (Hudson, Souriau 2003). Lipids and PEGs have also been coupled to antibodies to improve the stability of the molecule.

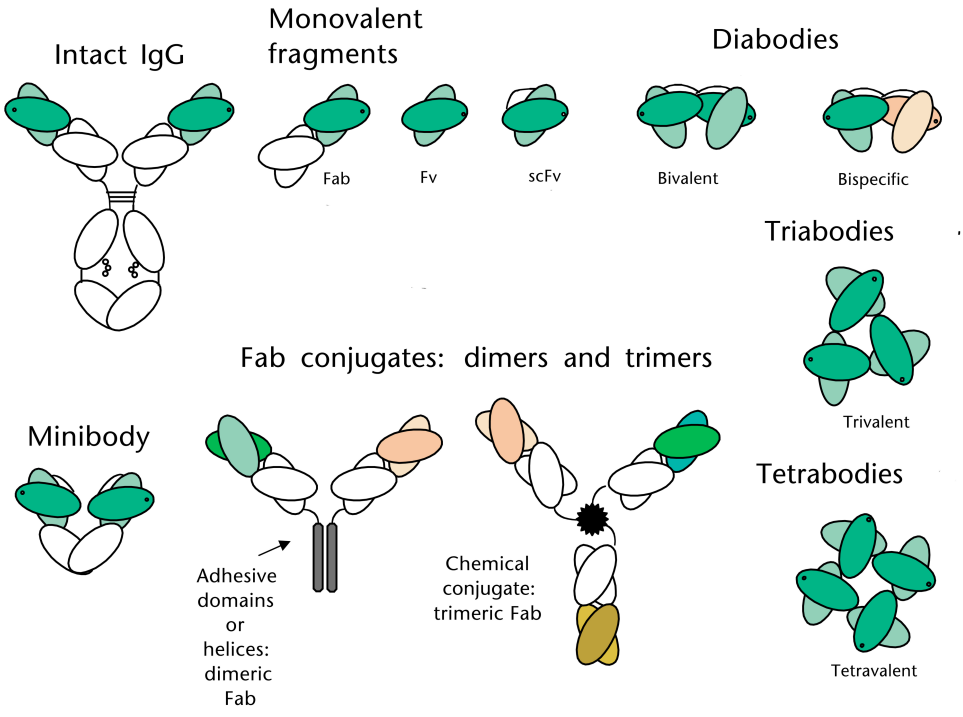


Figure 2. Schematic representation of an intact antibody (IgG) together with Fab and Fv fragments and single variable (V) (colored ovals; dots represent antigen-binding sites) and constant (C) domains (uncolored ovals). Figure reprinted by permission from Macmillan Publisher Ltd: [Nature Medicine] Hudson and Souriau (2003).

Bispecific antibodies (bsAbs) contain two different binding specificities fused together (Segal, Weiner & Weiner 2001). In the simplest case, the specificities are directed toward the same target, increasing the avidity, but bispecific antibodies are also used to crosslink two different antigens. In a recent study, Glaser and coworkers produced stable IgG-like bispecific antibodies that targeted and crosslinked two TNF family receptors (TNF-related apoptosis-inducing ligand receptor 2 and lymphotoxin beta receptor) (Michaelson *et al.* 2009). Other experiments showed that therapeutic treatment with the developed bispecific antibody resulted in a reduction in tumor volume that was similar or greater than that induced by treatment with a combination of parental antibodies. These so-called "two-in-one" antibodies have also been generated from the well-known therapeutic antibody Herceptin. This molecule binds human epidermal growth factor receptor 2 (HER2) and was engineered to simultaneously bind vascular endothelial growth factor (VEGF) (Bostrom *et al.* 2009). By combining different specificities and multivalent antibodies, it is possible to create multispecific antibody molecules (Todorovska *et al.* 2001).

2.1.2.1 Therapeutic antibodies

Monoclonal antibodies are the best-selling class of biologics in the US market, reaching 16.9 billion dollars in sales in 2009 (Aggarwal 2010). Additionally, the average growth in sales has been 10% during the past few years. At the moment, the US Food and Drug Administration (FDA) has registered 31 monoclonal antibodies for marketing (Aggarwal 2010) (Table 2). Antibody therapies for cancer and inflammatory disorders have been the most successful products.

For cancer treatments, there are more monoclonal antibodies against circulating cancer cells than against solid tumors (Table 2). This bias is due to greater accessibility of circulating cells for the antibody. Antibodies fused with radioactive substances are widely used in applications such as cancer imaging and therapy. The balance between long dissociation rates at the target site and slow blood clearance is important because accumulation in the liver together with high radiation exposure for other tissues present severe side effects of the treatment.

Table 2. Therapeutic antibodies approved by the US Food and Drug Administration (FDA). The table is adapted from Hudson and Souriau (2003) and from http://www.immunologylink.com/FDA-APP-Abs.html (19.8.2011).

Product name[a]	Specificity	Product type	Indication	Year
Orthoclone OKT3	CD3	Mouse	Transplant rejection	1986
ReoPro	gpIIb/gpIIa	Chimeric Fab	Cardiovascular disease	1994
Rituxan	CD20	Chimeric	Non-Hodgkin lymphoma	1997
Zenapax	CD25	Humanized	Transplant rejection	1997
Remicade	TNF-α	Chimeric	Crohn's disease, rheumatoid arthritis	1998-1999
Simulect	CD25	Chimeric	Transplant rejection	1998
Synagis	RSV	Humanized	Respiratory syncytial virus	1998
Herceptin	Her-2	Humanized	Metastatic breast cancer	1998
Mylotarg	CD33	Humanized	Acute myeloid leukemia	2000
Campath	DC52	Humanized	Chronic lymphocytic leukemia	2001
Zevalin	CD20	Mouse	Non-Hodgkin lymphoma	2002
Humira	TNF-α	Human	Inflammatory diseases	2002
Bexxar	DC20	Murine	Non-Hodgkin lymphoma	2003
Xolair	IgE	Humanized	Severe allergic asthma	2003
Raptiva	CD11	Humanized	Autoimmune disease	2003
Avastin	VEGF	Humanized	Metastatic colorectal cancer, non-small cell lung cancer, metastatic breast cancer	2004
Tysabri	α4 subunit of α4β1	Humanized	Multiple sclerosis, Crohn's disease	2004
Erbitux	EGFR	Chimeric	Colorectal cancer, head and neck cancer	2004
Vectibix	EGFR	Human	Metastatic colorectal carcinoma	2006
Lucentis	VEGF-A	Humanized Fab	Wet macular degeneration	2006
Soliris	CD59	Humanized	Paroxysmal nocturnal hemoglobinuria	2007
Cimzia	TNF-α	Humanized Fab	Crohn's disease, rheumatoid arthritis	2008
Simponi	TNF-α	Human	Rheumatoid and psoriatic arthritis, active ankylosing spondylitis	2009
Ilaris	IL1b	Human	Inflammatory disease	2009
Stelara	IL-12/23	Human	Inflammatory disease	2009
Arzerra	CD20	Human	Chronic lymphocytic leukemia	2009
Actemra	Anti-IL-6R	Humanized	Autoimmune disease	2010
Benlysta	BLyS	Human	Autoimmune disease	2011
Yervoy	CTLA-4	Human	Melanoma	2011

[a] Product names are registered trademarks.

Classically, therapeutic monoclonal antibodies have been used in combination with chemotherapeutic drugs to treat cancer because in many cases monoclonal antibod-

ies treatment is not effective enough on its own. Many strategies have also been investigated to improve the efficiency of therapeutic antibodies, including enhancement of intrinsic Fc-linked effector functions by glycoengineering and the use of bispecific antibodies, polyclonal antibodies, and conjugates (Figure 3, Beck *et al.* 2010a).

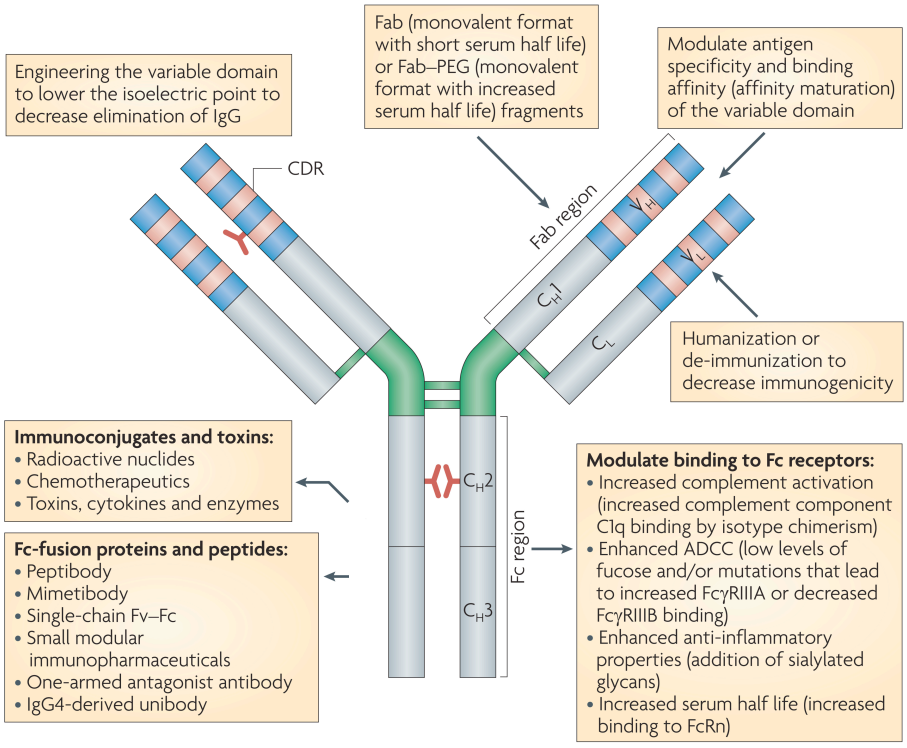


Figure 3. Antibody design to improve the pharmacological properties and therapeutic response. Figure reprinted by permission from Macmillan Publisher Ltd: [Nature Reviews Immunology] Beck *et al.* (2010b).

The key strategy of a new generation of antibody drug conjugates is to combine the cytotoxicity of natural or synthetic agents with monoclonal antibodies conjugated by optimized linkers (Beck *et al.* 2010a). However, the clinical success of these immuno-conjugates to date has been limited compared to unmodified antibodies because the development of antibody drug conjugates takes longer and is much more complex. The successful development of antibody conjugates depends on the optimization of antibody selection, linker stability, cytotoxic drug potency, and the mode of conjugation (Junutula *et al.* 2010). Gemtuzumab ozogamicin (Mylotarg) is a clinically approved antibody conjugated to calicheamicin that is used for the treatment of acute myeloid leukemia (see Table 2) (Hamann *et al.* 2002). This

therapeutic antibody targets and eliminates the CD33 receptor, which is highly expressed on cancer cells. Another successful example of a drug-conjugated antibody is trastuzumab emtansine (previously called T-DM1), which shows greater activity than the nonconjugated HER2-targeting antibody trastuzumab during HER2-positive breast cancer therapy (Lewis *et al.* 2008, Junutula *et al.* 2010)

The tumor-associated antigens that are over-expressed in B-cells, T-cells, carcinoma cells, endothelial, or stroma cells have been widely investigated (Beck *et al.* 2010a). Reducing the blood supply to tumors has been another efficient alternative treatment for cancer. Taken together, antibody therapeutics has been extensively studied during recent years, and third-generation antibodies having multiple simultaneous activities as discussed above, have entered clinical trials.

2.1.3 Challenges inherent to antibody usage

Despite the success of antibody molecules as therapeutic agents or as recognition molecules in diagnostic applications, there are considerable limitations to antibody usage. For instance, in in terms of drug targeting, conventional antibodies appear to have structural features that are less desirable including their rather large size and a composition based on two different protein chains (Skerra 2003). Large molecule size limits tissue penetration, and the presence of two different polypeptide chain structures can lead to unstable domain association. Complex structures also complicate the cloning steps that are required for recombinant expression and manipulation of the antigen-binding site of antibody. The stability of the antibody molecule can also be a limiting factor in some applications. Additionally, some antibody molecules tend to aggregate especially when fused to other molecules (Binz, Amstutz & Pluckthun 2005).

Despite the immense amount of different specific binding molecules that the antibody repertoire includes, some molecules are still not recognized by the immune system. Molecules with cavities or clefts, such as prions, have been refractory to conventional antibodies (Hudson, Souriau 2003). Interestingly, single V-like antibody domains with cavity-penetrating CDR -loops have been discovered in camelids (camels and llamas) (Desmyter *et al.* 2001) and in sharks (Nuttall *et al.* 2001). By tailoring these proteins (Vincke *et al.* 2009) or mimicking their structure in

in human antibodies, improved characteristics could be achieved to access immunosilent sites.

One severe drawback to antibody usage in human therapies is the ability to trigger a human anti-mouse antibody (HAMA) or human anti-rat antibody (HARA) response (Getts *et al.* 2010). Fully human antibodies have been utilized in many applications to overcome this drawback, but even these antibodies have been shown to induce marked immune responses (Getts *et al.* 2010, Harding *et al.* 2010). Surprisingly, in some cases, a humanized murine antibody may provide a therapeutic advantage when compared to fully human antibodies. Therefore, it is essential to identify and design predictive models to select appropriate therapeutic forms in the future (Getts *et al.* 2010).

The aspects that limit the industrial use of antibodies are the high cost of production and intellectual property issues. For example, the large-scale production of glycosylated full-sized antibodies requires time-consuming and expensive eukaryotic expression (Skerra 2007). To solve this problem recombinant non-glycosylated Fab and scFv fragments are efficiently expressed in bacterial expression systems with purification tags (e.g.,c-Myc, Histidine, FLAG). Affordable antibody treatments have been developed by increasing the productivity and using less costly downstream processing (Beck *et al.* 2010b). However, it must be taken into account that the intellectual property situation surrounding antibodies is quite complex and therefore, it can be difficult to protect research results (Skerra 2007).

As with other protein drug therapies, antibody cancer therapy may also lose its therapeutic efficacy due to acquired resistance to the antibody. Primary resistance toward therapeutic antibodies may also arise. For example, both of these drawbacks have been reported for the therapeutic use of the well-known antibody trastuzumab to treat HER2-positive breast cancer (Nahta *et al.* 2006; Kruser, Wheeler 2010).

2.2 Non-immunoglobulin affinity proteins provide a valuable alternative to antibodies

The development of life science technologies has made it possible to generate engineered non-immunoglobulin affinity proteins comparable to conventional antibodies in terms of specificity and affinity. Recently, protein engineering of non-

immunoglobulin affinity proteins has been extensively studied, and a number of reviews describing engineered scaffolds can be found elsewhere (Gebauer, Skerra 2009; Gronwall, Stahl 2009; Skerra 2007; Binz, Amstutz & Pluckthun 2005; Nygren, Skerra 2004; Hey *et al.* 2005; Hosse, Rothe & Power 2006). Approximately 50 different non-Ig protein scaffolds have been successfully modified (Gronwall, Stahl 2009). Examples of modified scaffolds and the protein engineering and selection methods use to generate them are listed in Table 3.

Engineered proteins vary in structure and in function, but the choice of the scaffold is mostly dependent on the indented use. If affinity proteins are to be used as an alternative to antibodies, they should lack the inherent limitations of antibodies. The ideal scaffold has a robust, stable architecture, as it must provide the capability to fold as a functional protein despite extensive reshaping, even in cases of amino acid replacements or insertions and deletions (Skerra 2003). The ideal scaffold should also be composed of a single polypeptide chain. Additionally, a low degree of post-translational modification is also advantageous (Gronwall, Stahl 2009; Skerra 2003). If the intended use of the engineered protein is therapeutic, the protein must also have appropriate pharmaceutical properties, such as a long serum half-life, and it should remain homogenous despite large-scale manufacturing or long-term storage.

Table 3. Examples of affinity proteins tailored from immunoglobulins and from alternative scaffolds.

Name	Scaffold	Randomization	Selection method	Antigen/ Target	Affinity (K_d)	Reference
Antibody	Fab	CDR-L3, -H1, -H2: Y,S CDR-H3: Y,S,G/Y,S,R/Y,S,G,R	Phage display	Insulin, VEGF and HER	0.3 – 188 nM	Birtalan *et al.* (2008)
Antibody	scFv	CDR1, CDR2, CDR3	Ribosome display	Bovine insulin	82 pM	Hanes *et al.* (2000)
Antibody	Fab	CDR-L3 CDR-H3	Phage display	Testosterone	0.3 nM	Hemminki *et al.* (1998a) Hemminki *et al.* (1998b)
Adnectin	Fibronectin	3 loops	mRNA display	TNF-α	20 pM	Xu *et al.* (2002)
Affibody	Protein A	13 amino acids on 2 α-helices	Phage display	ErbB2	22 pM	Orlova *et al.* (2006)
Anticalin	Lipocalin (BBP)	4 loops (16 amino acids)	Phage display	Digoxigenin, digoxin, fluorescein	30 nM 800 pM 1 nM	Beste *et al.* (1999) Schlehuber, Beste & Skerra (2000)
	Lipocalin (Lcn2)	4 loops (20 amino acids) by error-prone PCR	Phage display	CLTA-4	240 pM	Schonfeld *et al.* (2009)
DARPin	Ankyrin repeat	7 amino acids on a β-turn and 1 α-helix of every repeat	Ribosome display	Maltose binding protein, JNK2, p38	4.4-2.1 nM	Binz *et al.* (2004)
Kunitz domain	APP1	1 to 2 loops	Phage display	Human plasma kallikrein	1.2 nM	Dennis and Lazarus (1994) Williams and Baird (2003)

There are already a number of different biotechnological and medical applications available in which novel affinity scaffolds have been utilized (Figure 4). Engineered affinity proteins have been used as capture agents in affinity matrices when large amounts of protein are needed. The use of such proteins has been advantageous because they can be efficiently produced in prokaryotic expression systems and inexpensively purified. Protein A-derived affibodies have been successfully used as capture agents (Nygren 2008). These multipurpose molecules will be discussed in more detail in Section 2.2.2.

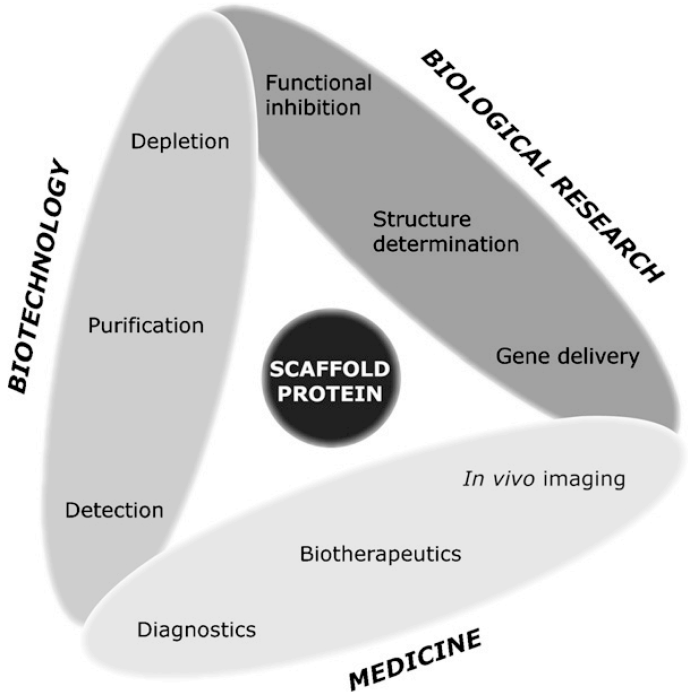


Figure 4. Possible applications of novel affinity scaffolds in biotechnology and medical applications. Reprinted from Gronwall and Stahl (2009) with permission from Elsevier.

Novel affinity molecules can also be used as detection reagents in traditional antibody-based procedures such as ELISAs, flow cytometry and immunohistochemistry. Modified affinity proteins such as affibodies (Hogbom *et al.* 2003) and DARPins (Binz *et al.* 2004) have been used for co-crystallization to determine the structures of proteins that are otherwise not possible to crystallize. Engineered proteins are highly promising as imaging agents for *in vivo* diagnostics (Gronwall, Stahl 2009). They can be used to identify or localize pathological conditions, as high af-

finity and specificity are essential requirements for proteins used for imaging. Small size and rapid body clearance are of great importance for imaging molecules.

Engineered proteins are also used as pharmaceuticals for biotherapeutic applications, and today, the majority of biopharmaceuticals are so called “protein-drugs” (Walsh 2005). Protein-drugs can interfere with or block functions, target molecules or organisms, or even stimulate a signaling pathway. Additionally, these proteins can serve as delivery vehicles.

2.2.1 Anticalins are artificial affinity proteins tailored from lipocalin structures

Anticalins are artificial binding proteins tailored from the plastic structures of lipocalins (Skerra 2008). Lipocalins are members of a protein family that includes several hundred different proteins, of which many are well characterized and have determined 3D -structures (Skerra 2008). Anticalins can bind small hydrophobic molecules and specific cell-surface receptors as well as form complexes with other soluble macromolecules (Flower, North & Sansom 2000). Lipocalins are used for the transportation or storage of organic compounds, especially vitamins, steroids, and lipids (Skerra 2008).

Lipocalins serve as an excellent scaffold for protein engineering; they have a small and simple molecular architecture (being 160 to 180 amino acid residues in length), and they are often thermostable. The lipocalin fold is a β-barrel structure with eight antiparallel β-sheets. The ligand binding site is composed of an internal cavity and an external loop scaffold (Flower 1996). The absence of post-translational modifications is also advantageous for a modified protein scaffold. The robust structure of lipocalins tolerates structural modifications and mutations quite well. Additionally, the ligand-binding sites of lipocalins resemble the CDRs of antibodies; therefore, this site has considerable potential for molecular engineering and recognition technology.

Novel affinity proteins derived from the lipocalin fold have been successfully selected using small hapten-like compounds and large protein antigens. Additionally, lipocalins have been extensively modified, and even forms where two lipocalins have been fused together to form duocalin have been generated (Schlehuber, Skerra 2001). Because the diversity of the different lipocalins cannot be acquired by so-

matic gene recombination and hypermutation, as in the case of antibodies, novel methods of combinatorial biochemistry have been developed and applied to engineer lipocalins as novel affinity proteins. For instance, the lipocalin from *Pieris brassicae*, bilin-binding protein (BBP), has been successfully engineered to recognize novel ligands (Beste *et al.* 1999), such as fluorescein (Gotz *et al.* 2002) and digoxigenin (Schlehuber, Beste & Skerra 2000). Importantly, this modification leads to affinities comparable to those of antibodies (fluorescein K_D ~ 1 nM; digoxigenin K_D = 30 nM). Possible applications for a fluorescein-binding anticalin may include use as a reagent for the specific quenching of background signals in biophysics studies. A dioxigenin-binding anticalin was subjected to affinity maturation by a combinatorial approach of loop-walking randomization and rational protein design. This successful strategy led to the capture of a high-affinity digoxin binding protein with a K_D value 800 pM (digoxin is a natural glycosylated derivative of digoxigenin) (Schlehuber, Skerra 2002). This novel molecule may be applicable as a therapeutic agent for the treatment of digitalis intoxication.

Human anticalin libraries have been created and selected against a variety of disease-related protein antigens. For example, an antagonist with an inhibitory effect on the cytotoxic T-lymphocyte antigen-4 (CTLA-4) receptor was found using a neutrophil lipocalin library. This protein may be used as a drug in the immunotherapy of cancer (Skerra 2008). An anticalin recognizing vascular endothelial growth factor (VEGF) was also foud; this molecule exhibits favorable binding when compared to currently approved VEGF antagonists (Skerra 2008). Human lipocalins can be used as carrier molecules or scavengers for pharmaceutically active compounds because they are soluble in plasma and also in other body fluids in quite high concentrations (up to 1 mg/mL) (Skerra 2008).

One challenge for the use of lipocalins is their rapid clearance by renal filtration. This clearance is disadvantage in many therapeutic applications. To extend the serum half-life, anticalins can be fused with serum albumin or with an albumin-binding domain (Skerra 2008). Alternatively, the addition of polyethylene glycol (PEG) to the anticalin molecule has been used for the same purpose. The properties and possible application of anticalins have been widely reviewed elsewhere (Beste *et al.* 1999; Schlehuber, Beste & Skerra 2000; Skerra 2008; Schlehuber, Skerra 2002; Skerra 2001).

Pieris AG is a biotechnology company employing Anticalin® technology to develop more efficacious and convenient protein therapeutics (http://www.pieris-ag.com/index.php (18.6.2011)). The company was founded in 2001 by Dr. Arne Skerra and Claus Schalper.

2.2.2 Affibodies are affinity proteins derived from *Staphylococcal* protein A

Affibodies are the ultimate example of artificial binding molecules. These affinity proteins have been selected toward a large number of targets and used in several applications, such as bioseparation, diagnostics, functional inhibition, viral targeting, and *in vivo* tumor imaging and therapy (Nygren 2008). *Staphylococcal* protein A-derived affinity molecules are ideal scaffolds for protein engineering (Figure 5). Affibodies are small (6 kDa), cysteine-free molecules with a structure containing a three-helix bundle domain. They are easy to produce in soluble and proteolytically stable forms, and their robust structure is highly tolerant of modifications and fusion to other protein or peptides.

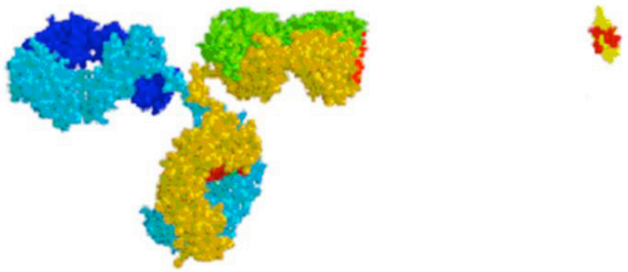


Figure 5. Comparison of the structures of monoclonal antibody and affibody molecule. Affibody (6 kDa) is significantly smaller and simpler in structure than monoclonal antibody (150 kDa). Figure adapted from http://www.affibody.com/en/Product-Portfolio/Technology/Affibody-Molecules---Technology/ (18.6.2011).

Originally, protein A molecules were employed in the purification and detection of antibodies. However, due to a specific interaction with antibodies, protein A was subjected to mutagenesis to change its ligand recognition to target other molecules. A gene library based on protein A, in which 13 amino acids from helices 1 and 2 were randomly mutated was created and tested for novel binding proteins via phage

display selection (Nord *et al.* 1997). Since then, specific affibody molecules have been generated toward various ligands including human epidermal growth factor receptor 2 (HER2) (Wikman *et al.* 2004), HIV-1 envelope glycoprotein gp120 (Wikman *et al.* 2006), and interleukin-2 receptor α (Gronwall *et al.* 2008). Remarkably, generated binders provide affinities comparable to those obtained for naïve libraries of antibody fragments.

In addition to selection for novel binding properties, affibodies have been fused to several different molecules. For example, antibody-like fusion constructs have been generated and successfully expressed in *E. coli*, where the Fab-arms were replaced with affibodies (Ronnmark *et al.* 2002). These artificial antibodies may represent alternatives for antibodies in diagnostic or therapeutic applications. Additionally, affibodies have been fused to various labels for detection in techniques like flow cytometry, immunofluorescence and immunohistochemistry. Technetium-labeled HER2-targeting affibodies have been used for radionuclide tumor targeting *in vitro* and *in vivo* (Orlova *et al.* 2006). Affibody molecules have been used also to re-direct viral vectors to cancer cellsas part of gene therapy (Henning *et al.* 2002). This technique was taken further, and an adenovirus vector was engineered to recognize two different ligands, HER2 and *Taq* polymerase, for use in gene transfer assays (Myhre *et al.* 2009). Furthermore, there are commercially available applications based on affibody technology, such as a DNA polymerase-binding affibody that creates hot-start PCR conditions when added to the reaction (Phusion Hot Start High-fidelity DNA Polymerase, Finnzymes, Thermo Fisher Scientific).

The Swedish biotech company called Affibody (http://www.affibody.com/en/ (18.6.2011)), launched in 1998, focuses on developing biopharmaceuticals based on Affibody® molecules.

2.2.3 The avidin scaffold has potential as a novel affinity protein

Avidin is a tetrameric protein found in the chicken egg white, and it is originally known for its capacity to bind biotin with a high-affinity (Eakin, Snell & Williams 1941). Like lipocalins, avidins have beta-barrel structures, and both belong to the calycin superfamily (Flower 1993). Due to its extreme stability and its ultra-tight ligand binding (K_d for biotin ~10^{-15} M (Green 1975)), avidin provides an attractive

scaffold for protein engineering. Avidin proteins have been extensively modified during recent years (Laitinen *et al.* 2006) and utilized in many biotechnology applications (Laitinen *et al.* 2007).

The structure of avidin is well–known, and the 3D -structure has been determined (Livnah *et al.* 1993b, Pugliese *et al.* 1993, Repo *et al.* 2006), enabling the use of sophisticated mutagenesis methods. Avidin consists of four identical subunits, and each subunit consists of 128 amino acid residues (DeLange, Huang 1971). The avidin protein is basic, unlike its bacterial analog streptavidin, which is neutral. It is important to remember that avidin is a glycoprotein and has one intramonomeric disulfide bridge when considering its use for protein engineering. However, despite its post-translational modifications, avidin can be expressed efficiently in *E. coli* (Hytönen *et al.* 2004).

We have used avidin as a scaffold for protein engineering and modified the binding properties of avidin by targeted random mutagenesis. An avidin form (sbAvd-2) that binds steroids instead of its natural ligand was generated (Riihimaki *et al.* 2011). In the study, the loop area of the ligand binding site was randomized, and the phage display method (Smith 1985) was used for selection. We have also created other avidin mutants that bind small molecules that are considerably different from biotin (Hiltunen S. *et al.*, unpublished results).

The structure of avidin has been modified extensively during the past few years. For instance, two circularly permutated avidin subunits have been fused together to form dual-chain avidin (Nordlund *et al.* 2004, Nordlund *et al.* 2005). To improve properties of dual chain avidin, a chimeric dual-chain avidin with enhanced thermal stability and expression in *E. coli* was generated (Riihimaki *et al.* 2011).

2.3 Generation of novel affinity proteins by affinity selection from gene libraries

Recognition molecules are required for various life-science applications. Protein engineering of alternative protein scaffolds addresses this challenge by complementing traditional antibody-based biotechnology applications. Currently, novel affinity proteins with desired characteristics can be acquired by combinatorial protein engineering and library selection (Gronwall, Stahl 2009; Binz, Amstutz & Pluckthun

2005). The solubility, stability, resistance to aggregation, and expression level of the protein can also be notably improved with these techniques (Binz, Amstutz & Pluckthun 2005). The challenge, however, is to predict the changes needed to alter the protein function but still retain its structure. Fortunately, in many cases, an affinity protein can be changed significantly without changing the basics of the protein-ligand interface (Reichmann *et al.* 2007).

As discussed in Section 2.1.1.2 to acquire the vast diversity of naturally occurring antibodies, mammals use somatic hypermutation and selection (Figure 6B). This cyclic process is mimicked in the methods of protein engineering to produce high-affinity proteins (Skerra 2003). So far, the most successful methods for modifing proteins rely on several cycles of mutation, display, selection and gene amplification (Hudson, Souriau 2003) (Figure 6A). These cycles of mutation and selection can be carried out both *in vivo* and *in vitro*.

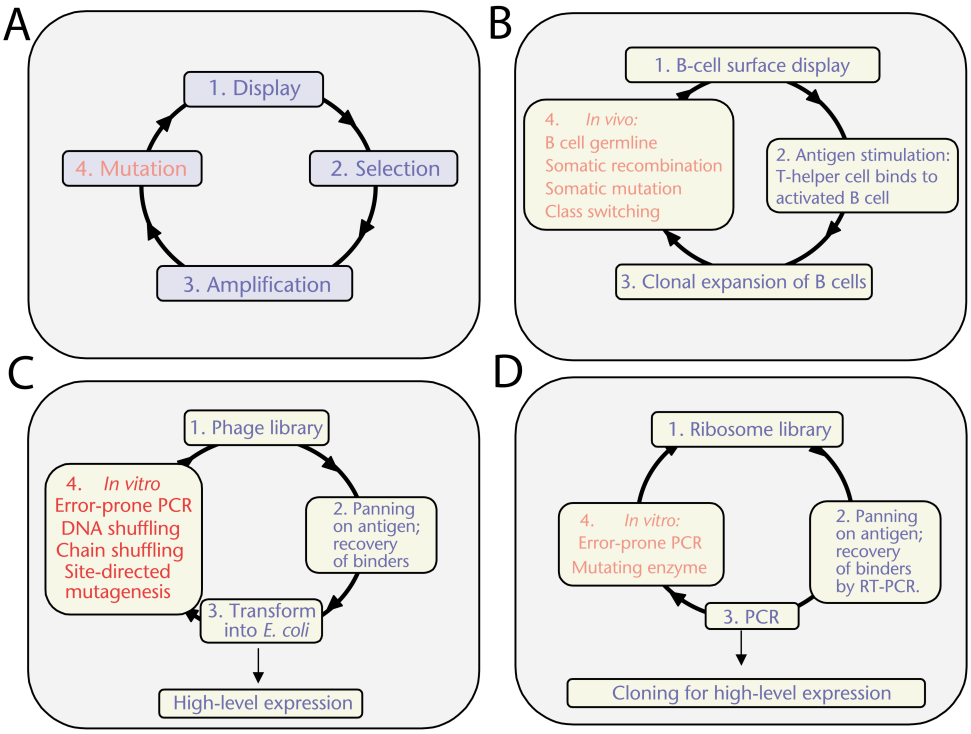


Figure 6. Selection methods rely on the several cycles of mutation, display, selection and gene amplification. Phage display (C) and ribosome display methods (D) are described more closely in Sections 2.3.2.1 and 2.3.2.2 respectively. Figure reprinted by permission from Macmillan Publisher Ltd: [Nature Medicine] Hudson and Souriau (2003).

2.3.1 Mutagenesis strategies for affinity protein modification

Because tailored proteins vary in their structures and functions, and the requirements for affinity proteins are diverse, there are a variety of different ways to modify the genes encoding such proteins. If the 3D -structure of a protein has been solved, rational design (Figure 7H), in which mutations can be targeted to known spots or regions of the gene, can be applified. This technique requires specific knowledge about the ligand-protein interaction. Detailed information about the characteristics of the amino acid residues that are the building blocks of proteins is also needed. Alanine scanning, in which each corresponding amino acid is changed to alanine, is often used to detect the effects of certain amino acid residues on the function of a protein (DeLano 2002).

With extensive knowledge of the ligand-binding site, an affinity protein with high affinity and specificity can be created. The shape complementarity between protein and ligand is primarily generated by aromatic side chains and specific interactions with suitably placed hydrogen-bond donors and acceptors (Skerra 2008). Computational methods are advancing rapidly, and as a result, computationally designed novel proteins with improved protein stability and functionality have been reported (Saven 2011). Hellinga *et al.* presented a computational method whit they constructed soluble receptors with high affinity and specificity for different ligands (Looger *et al.* 2003).

In some rational engineering approaches, a pre-existing binding domain has been grafted onto a recombinant protein (Figure 7G). For example, Desmadril *et al.* transferred the CDR3 of the VHH chain of the camel anti-lysozyme antibody to the structure of the small antibody-like protein neocarzinostatin (Nicaise *et al.* 2004). The resulting molecule was stable and specifically bound lysozyme.

However, the rational design of an affinity protein can be challenging because different proteins have diverse solutions to accomplish the requirements of specificity and affinity binding. Therefore, in many cases, combinatorial engineering is used instead of rational mutagenesis. In combinatorial mutagenesis the sequences of specific positions can be diversified by randomized codons. To modify or change the affinity of the protein, the randomization is generally targeted to the ligand-binding site, for instance to loop areas (Figure 7A) or to the ligand cavity (Figure 7D). These strategies are discussed in more detail in Section 2.3.1.1.

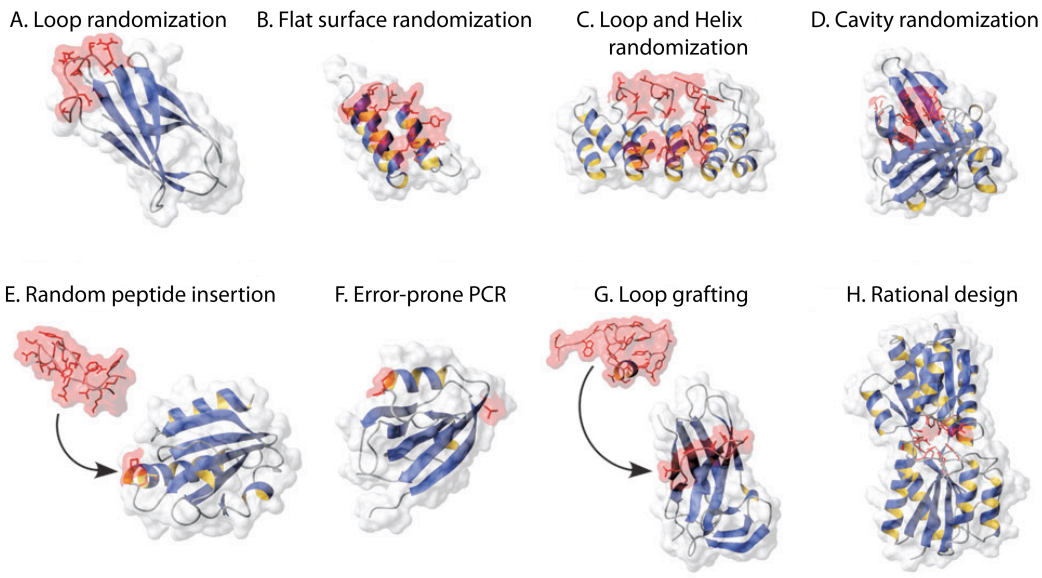


Figure 7. Mutagenesis strategies frequently used in affinity protein engineering. Figure reprinted by permission from Macmillan Publisher Ltd: [Nature Biotechnology] Binz, Amstutz & Pluckthun (2005).

Combinatorial strategies include random peptide insertion (Figure 7E) and error-prone PCR methods (Figure 7F). In a study by Brent *et al.*, the random peptide insertion strategy was used, and 20-residue peptide aptamers were transferred into the loop area of *E. coli* thioredoxin (Colas *et al.* 1996). Affinity selection of the aptamers resulted in a molecule that bound cyclin-dependent kinase 2. When the structure of a protein is unknown, the error-prone PCR method (Cadwell, Joyce 1992) may be a good choice for a mutagenesis strategy. In this method, random point mutations are introduced throughout the sequence due to the decreased fidelity of the DNA polymerase used in the PCR reaction. For instance, PDZ domains have been modified to specifically bind novel ligands using this method (Schneider *et al.* 1999).

One commonly used strategy is combining mutagenesis methods. Skerra *et al.* generated a molecule recognizing the T-cell core receptor CLTA-4 with subnanomolar activity from human lipocalin Lcn 2 using a combination of targeted mutagenesis and error-prone PCR (Schonfeld *et al.* 2009). In that study the loop area of lipocalin (20 amino acid residues) was randomized, and a library of 2×10^{10} individual molecules was created. To increase the affinity and specificity, molecules were affinity-matured by *in vitro* evolution using an error-prone PCR. As a result, Skerra

and colleagues captured an engineered lipocalin with an affinity (240 pM) approximately 10-fold better than the affinity of a humanized CTLA-4-specific antibody that is under clinical investigation (Phan *et al.* 2003).

Finally, *in vitro* homologous recombination is a good alternative to engineer proteins when the function or the structure of the protein is not well-known. For example, enzymes are difficult to engineer using rational mutagenesis, but the functions of enzymes have been improved successfully with homologous recombination *in vitro* (Minshull, Stemmer 1999).

2.3.1.1 Oligonucleotide-directed mutagenesis

A selective random mutagenesis strategy based on randomized codons is frequently used when creating affinity protein gene libraries. In the development of novel affinity scaffolds, strategies in which mutagenesis is targeted to a loop area (Beste *et al.* 1999; Schonfeld *et al.* 2009; Dennis, Lazarus 1994), flat surfaces (Nord *et al.* 1997), combinations of loops and helices (Binz *et al.* 2004), or cavities (Beste *et al.* 1999) have been reported (Binz, Amstutz & Pluckthun 2005).

Library design requires extensive planning because proteins may contain thousands of amino acid residues, and they may fold into various conformations. This property results in a large configurational space (Saven 2011) as well as in a large sequence space. Thus, it is challenging to design affinity proteins with well-folded structures and proper functions because the function of a protein is dependent on a well-ordered structure. In general, randomly generated sequences often fail to fold into proper conformations (Hecht *et al.* 2004).

Szostak *et al.* hypothesized that functional proteins are sufficiently common (roughly 1 in 10^{11}) that they can be discovered from the protein sequence space (Keefe, Szostak 2001). To address this hypothesis, they generated an enormous library of 6×10^{12} individual proteins containing contiguous stretches of 80 random amino acids and successfully selected functional ATP-binding proteins from the library.

Due to practical considerations, in many cases, six to eight amino acid residues at maximum can be totally randomized in a library (Clackson, Wells 1994). In the case of eight amino acids, there would be 20^8 different types of proteins in the library

because there are 20 naturally occurring amino acids. The library size can be decreased considerably using degenerate codons (Table 4). One codon commonly used in library design is NNK (Barbas *et al.* 2001, Cwirla *et al.* 1990), in which N is any nucleotide and K is either guanine or thymine. This codon allows for all 20 amino acids to be present. This set includes one stop codon instead of three possible stop codons. In addition, the number of codons is decreased from 64 to 32, which makes it easier to cover the sequence space.

Table 4. A set of degenerate codons used frequently in library design. The table is taken from Tonikian *et al.* (2007).

Codon	Description	Amino acids	Stop codons	Unique codons
NNK	All 20 amino acids	All 20	TAG	32
NNC	15 amino acids	A,C,D,F,G,H,I,L,N,P,R,S,T,V,Y	None	16
NWW	Charged, hydrophobic	D,E,F,H,I,K,L,N,Q,V,Y	TAA	16
RVK	Charged, hydrophilic	A,D,E,G,H,K,N,R,S,T	None	12
DVT	Hydrophilic	A,C,D,G,N,S,T,Y	None	9
NVT	Charged, hydrophilic	C,D,G,H,N,P,R,S,T,Y	None	12
NNT	Mixed	A,D,G,H,I,L,N,P,R,S,T,V,F,I,L,V	None	16
VVC	Hydrophobic	A,D,G,H,N,P,R,S,T	None	9
NTT	Hydrophobic	F,I,L,V	None	4
RST	Small side chains	A,G,S,T	None	4
TDK	Hydrophobic	C,F,L,W,Y	TAG	6

Sequence symbols: A (adenine), C (cytosine), G (guanine), T (thymine)
Wobble IUPAC-IUB symbols: N (A/C/G/T), K (G/T), W (A/T), V (A/C/G), R (A/G), S (C/G), D (A/G/T)

To enhance the discovery of functional proteins from the sequence space, features of rational design must be used to focus sequence diversity into regions that are most likely to yield proper structures (Hecht *et al.* 2004). Amino acid residues have specific roles based on their different structures. For instance, cysteine residues increase stability in small proteins by forming disulfide bridges, whereas proline residues, abundant in α-helical domains, are known to induce kinks in the protein structure. The forces that guide a protein to its folded form are largely based on the non-covalent interactions among of amino acid side chains, such as van der Waals, hydrophobic, electrostatic, and hydrogen-bonding interactions.

Limiting the sequence space of the library in favor of those amino acids that are most likely to contribute to affinity and/or specificity is beneficial, especially when aiming to alter or change the affinity of a protein (Birtalan *et al.* 2008). Some amino acids are over represented in ligand-recognition sites. For example, in a study of 50

different proteins (including streptavidin), glycine, serine, arginine and tyrosine were the most abundant residues found 4 Å from the ligand (Villar, Kauvar 1994). As mentioned in Section 2.1.1.1, serine and tyrosine were found to also be highly abundant in the ligand-binding sites of antibodies. Amino acid residues such as glycine, serine, arginine and tyrosine are often found in β-turn regions, where they have considerable influence on the characteristics of the ligand-recognition site. Due to the different structures of amino acid residues, interactions with the ligand are based on distinct mechanisms; serine and tyrosine form hydrogen bonds, positively charged arginine residues form salt bridges, and aromatic tryptophan moieties create hydrophobic interactions (π-π stacking) with the ligand. Minimalistic libraries that include only two amino acid residues in the ligand-binding site have been generated: Sidhu *et al.* obtained functional and high-affinity (K_D=60 nM) antibodies from a minimalistic library in which the amino acid residues of an antigen-binding site were changed to tyrosine or serine (Fellouse *et al.* 2005).

In an interesting study by Hecht *et al.*, a binary patterning strategy was used in the design of *de novo* four-helix bundle proteins (Kamtekar *et al.* 1993). In this strategy, hydrophobic and hydrophilic residues were explicitly selected and placed, as shown in figure 8. This library design was found to be very efficient, as half of the library members showed an affinity for heme (Rojas *et al.* 1997).

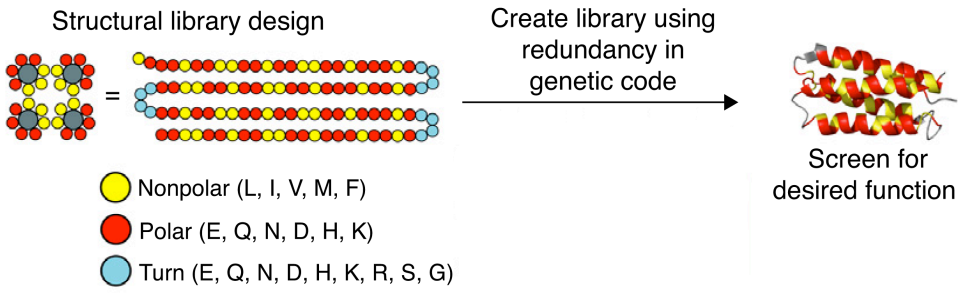


Figure 8. Strategy of four-helix bundle library using binary patterning. Nonpolar amino acids are placed in the ligand-binding cavity, whereas polar amino acids are located in the surface area. By extensive planning of codon usage, the sequence space of library can be limited and as a result the selection will be more efficient. Reprinted from Smith and Hecht (2011) with permission from Elsevier.

2.3.1.2 Homologous recombination in vitro

Homologous recombination *in vitro* is an effective mutagenesis method, that mimics sexual recombination in the evolution of new traits. During recombination, parental genes are combined together to create novel characteristics. This strategy generates enormous libraries; therefore, the major challenge is to identify the functional mutations from the background of neutral mutations (Giver, Arnold 1998). However, the method presents mutants throughout the entire target sequence which is not possible with many other strategies.

In 1994, Stemmer developed a DNA shuffling method (Stemmer 1994a, Stemmer 1994b) that enabled the *in vitro* formation of recombinant genes from parental genes. In this method, the parental genes were randomly fragmented by the DNase I –enzyme, and the purified fragments were then extended by repeated cycles of overlap extension into full-length genes.

Throughout the years, other methods such as random-priming *in vitro* recombination (RPR) (Shao *et al.* 1998) or staggered expression process (StEP) (Zhao *et al.* 1998) have also been developed. These methods are based on the *de novo* synthesis of template products and therefore require a much smaller amount of template. The DNase I fragmentation step is no longer needed.

2.3.1.3 Challenges in library design and construction

Introduced mutations are not always predictable. A distinctive drop in binding affinity can arise if destabilizing events such as the loss of van der Waals contacts, electrostatic pairings or buried nonpolar surfaces occurs. Protein aggregation or complete unfolding will also lead to a distinctive drop in binding affinity (DeLano 2002).

Library design can easily lead to biases. As discussed in Section 2.3.1.2, codons can be selected to limit the library size through the use of codons that allow for the insertion of only certain amino acid residues. This selection will cause controlled bias in the library. Unfortunately, gene libraries may be biased in unfavorable ways. The degenerate codon motif NNK (Table 4) is frequently used codon in random mutagenesis. However, the use of the NNK codon results in a large bias toward amino acids having more than one possible codon. There is also a significant influence on

ence on the part of the expression host selected for library generation because different organisms prefer certain codons over others (Sharp *et al.* 1988). In the error-prone PCR method, the genetic material is biased toward AT→GC transitions and AT→TA transversions. This effect results in a strong bias toward dAMP and dTMP substitutions, making these nucleotides more likely to be mutated (Vanhercke *et al.* 2005).

During transformation of the plasmids into bacterial cells, plasmids containing short inserts or no insert can allow for the overgrowth of bacterial clones containing them (Barbas *et al.* 2001). Furthermore, these clones can be overrepresented during library amplification, generating an unfavorable bias. If stop codons are allowed in the mutagenesis strategy, this choice can result in inserts that are too short to be expressed as functional domains, again giving a competitive growth advantage to bacteria harboring these clones. However, stop codons can be avoided using degenerate codon motifs (Table 4).

A sufficient library size is essential for the successful screening of novel binding proteins. At a minimum, the functional library size should exceed the sequence space of the library. In a functional library, only correctly assembled clones without any frameshift, stop codon or deletion will contribute to diversity (Knappik *et al.* 2000). Of note, this number is usually well below the reported diversity obtained by counting the number of transformants.

2.3.2 Selection systems

Selection systems are needed to screen gene libraries for proteins with novel, desired properties. A number of different selection methods have been developed, because the selection conditions may vary significantly from one protein to another, as the selection depends on the nature of the tailored protein. For instance, phage display is a powerful method for peptides and small, simple proteins, but if an affinity protein is complex in structure and includes disulfide bridges, the protein may not be folded correctly during the phage assembly process inside a bacterial cell. In such cases, another selection system, e.g., eukaryotic yeast display, would be more efficient. Selection systems can be divided into three categories: cell-dependent systems, cell-free systems, and non-display systems relying on the protein interaction

between an affinity protein and its target to allow for survival or generate fluorescence (Gronwall, Stahl 2009).

In cell-dependent systems, the affinity proteins are displayed on the surface of phage or cells. The most frequently used method is phage display (Smith 1985), in which an affinity protein is displayed on a bacteriophage, but due to the development of flow cytometry techniques, bacterial display (Daugherty 2007) has also become more popular. Libraries with sizes up to 10^{13} can be created with cell-dependent methods. However, the use of these systems is limited by the need for DNA transformation and by the limits of *in vitro* transcription and translation (Gronwall, Stahl 2009).

When larger gene library sizes are desired, methods that utilize *in vivo* transcription and translation should be used to avoid these limitations. Cell-free systems such as the ribosome display (Mattheakis, Bhatt & Dower 1994) and mRNA display (Weng *et al.* 2002) enable the creation of enormous libraries, and these methods are more amenable to automation. The mutagenesis step can be performed during selection. Strategies such as homologous recombination *in vitro* produce enormous gene libraries (Giver, Arnold 1998), and cell-free systems are appropriate choices to screen such libraries.

2.3.2.1 Phage display

Phage display is based on bacteriophages, viruses that infect bacteria, and was initially created for the affinity selection of protein fragments. The method was first described by Smith in 1985, and since that time, the method has been used successfully to generate several different affinity peptides and proteins (Clackson, Wells 1994; Winter *et al.* 1994; Hoogenboom *et al.* 1998; Hoogenboom, Chames 2000; Uchiyama *et al.* 2005). In this method, viruses are used to link protein recognition to DNA replication by displaying the protein on the surface of the phage that carries the corresponding gene in its genome. Large gene libraries can be cloned into phages, thus creating a powerful tool to analyze and select proteins with desired properties. Based on the properties of the displayed protein, an individual phage can be captured and amplified through the infection of bacteria. The new generation of

phages is then used for another selection round. The principle of this method is shown in figure 9.

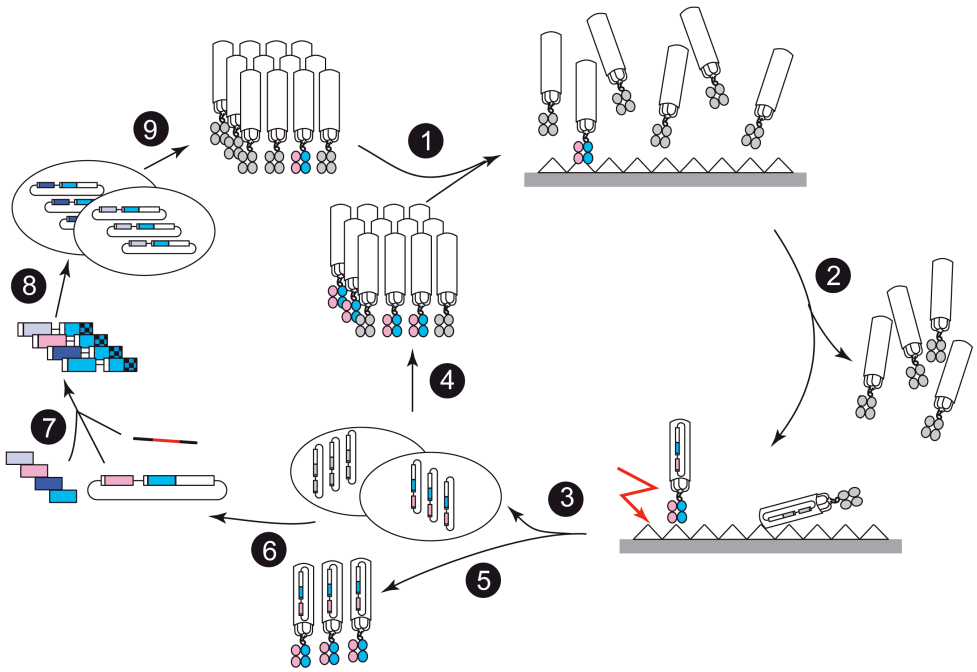


Figure 9. The principle of phage display and affinity maturation. A diverse phage library is incubated with the target molecules (1) that are immobilized for example on a solid surface. Washing steps remove the non-specifically bound phages (2) and the binders are then released typically by using a pH shock (3). Binders are then amplified by infection and multiplication in *E. coli* (4) to generate a new generation of highly enriched phages with binding sites specific for the antigen. This cycle is repeated two to four times until desired binders dominate the population (5). To affinity-mature selected binders, their genes are isolated from the display vector (6) and diversity is introduced by mutagenesis (7). After mutagenesis genes are sub-cloned (8) and phage particles of the new repertoire produced (9), to be used in re-selection procedure (1–4) with increasing stringency. Figure reprinted from Hoogenboom and Chames (2000) with permission from Elsevier.

Different types of phages, such as lambda (Huse *et al.* 1989), have been used for the phage display method, the Ff class of filamentous phages (f1, fd, and M13) being the most extensively studied (Barbas *et al.* 2001). Ff -viruses use the F conjugative pilus as a receptor, and therefore they are specific for *E. coli* carrying the F plasmid. Importantly, infection by the Ff phage is well tolerated by the host, causing only a 50 % longer generation time than that of uninfected bacteria.

During assembly, circular single-stranded DNA is packaged within the protein capsid cylinder of the bacterial envelope. The protein capsid cylinder consists of 11 different proteins (Figure 10). The length of the cylinder is dependent on the length of the DNA, leading to little constraint on the size of the DNA packaged (Barbas *et al.* 2001).

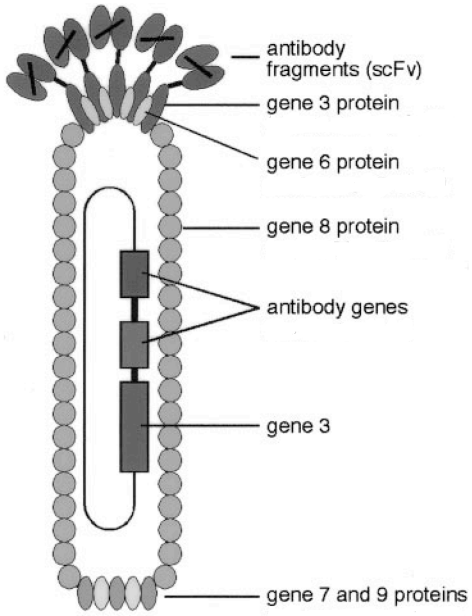


Figure 10. Structure of scFv displaying filamentous phage. Figure is from Smith *et al.* (2005).

Many phage coat proteins have been used as fusion partners in this method, but the most frequently utilized coat proteins, p8 and p3, are discussed here. The major coat protein p8 is the most abundant of the capsid proteins. There are approximately 2700 molecules of p8 in one virion (Barbas *et al.* 2001). Casareni and colleagues have studied the effect of fusing peptides to p8 on phage propagation, and they found that phage viability is affected by the length of the corresponding peptide and that peptide sequence plays a minor "tuning" role (Iannolo *et al.* 1995). At maximum, 6 to 8 amino acid residues can be added for functional display when peptides are fused to p8 (Iannolo *et al.* 1995). The reason for this limit is unknown, but the assembly of the phages may be interfered with by the fusion partner. In addition, inserts may disturb the rate of successful targeting and translocation of the chimeric p8 into the membrane (Barbas *et al.* 2001). Larger peptides or even proteins can be

displayed in a hybrid virion where 80 % of p8 proteins are wild type, and the remaining 20 % are fused p8 molecules (Greenwood, Willis & Perham 1991). The choice of p8 as a fusion partner instead of p3 can be based on two things: i) in some experiments, it is advantageous to have multiplicity in the binding sites along the phage surface, and (ii) the p8 fusion will not interfere with phage infectivity (Kang *et al.* 1991).

The most utilized capsid protein for the display of proteins is p3. Because there are only approximately five molecules of p3 per phage molecule, the chimeric p3 proteins containing large inserts are tolerated and packaged well into phage (Barbas *et al.* 2001). Wt p3 includes three domains that are separated by glycine-rich regions (Deng, Malik & Perham 1999). Domain 1 (amino acids 1-66) is required during infection for translocation of the phage DNA into the bacterial cytoplasm and for insertion of the coat proteins in to the membrane. Domain 2 (amino acids 86-216) is utilized when the phage binds to the F pilus during the initial phage of the infection. Domain 3 (amino acids 257-377) is essential for forming a stable phage particle by anchoring p3 to the tip of the virion. Because of the central role of the p3 protein in the phage infection process, the insertion of a large insert may lead to non-infective phages (Smith 1985). For monovalent display of larger proteins, hybrid phage-based method can be used. In these methods, a fusion of a novel protein to p3 is expressed from a phagemid, and wt p3 is encoded by a helper phage. Proteins may be introduced into the c-terminal part of p3 from which domains 1 and 2 are missing (Barbas *et al.* 2001, Barbas *et al.* 1991).

Selectively -infective phages, generally called as SIP technology, may be used for identifying robust and proteolysis-resistant proteins by phage display selection (Krebber *et al.* 1995, Krebber *et al.* 1997). The method may also be called selection and amplification of phage (SAP) (Duenas, Borrebaeck 1994) or direct interaction rescue (DIRE) (Gramatikoff, Georgiev & Schaffner 1994). The principle of this method is to select interacting protein-ligand pairs by fusing a ligand-binding protein with the N-terminal domains of p3 protein. The N-terminal domains of p3 will be then linked to the ligand. Infectivity of the phage particle is only restored when the displayed protein binds to the ligand. This property leads to the rapid selection of high-affinity interactions (Krebber *et al.* 1997).

2.3.2.2 Ribosome display

Ribosome display was first developed for peptide libraries (Mattheakis, Bhatt & Dower 1994). Ribosome display differs from phage display in that the ribosome complexes are constructed totally *in vitro*, and the capture of highly specific binding proteins is therefore faster (Hanes *et al.* 2000; Hanes, Pluckthun 1997). Furthermore, this method enables the generation of large gene libraries, up to even 10^{14} members, and during the selection process, additional mutations can be easily added to bring about affinity maturation. The ribosome display method is closely related to mRNA display, which is another cell-free selection system. In mRNA display, the modified protein is covalently bound to mRNA (Weng *et al.* 2002).

In this method, genotype and phenotype are linked through ribosomal complexes, which consist of mRNA, the ribosome, and the encoded protein. Protein is used for selection. In this *in vitro* selection method (for the principle of the method, see Figure 11), the first step is to amplify the DNA library by PCR, whereby a T7 promoter, ribosome binding site, and stem-loops are introduced and transcribed into RNA. RNA is then purified and translated *in vitro* in an *E. coli* S-30 system. Translation is stopped, and complexes are stabilized by cooling on ice and by increasing the magnesium concentration. The desired ribosome complexes are affinity selected by binding to an immobilized ligand. The bound ribosome complexes are dissociated by EDTA or with free antigen. Finally, RNA is isolated from the complexes and reverse transcribed to complementary DNA. The cDNA is then amplified by PCR and used for the next cycle of selection (Hanes, Pluckthun 1997).

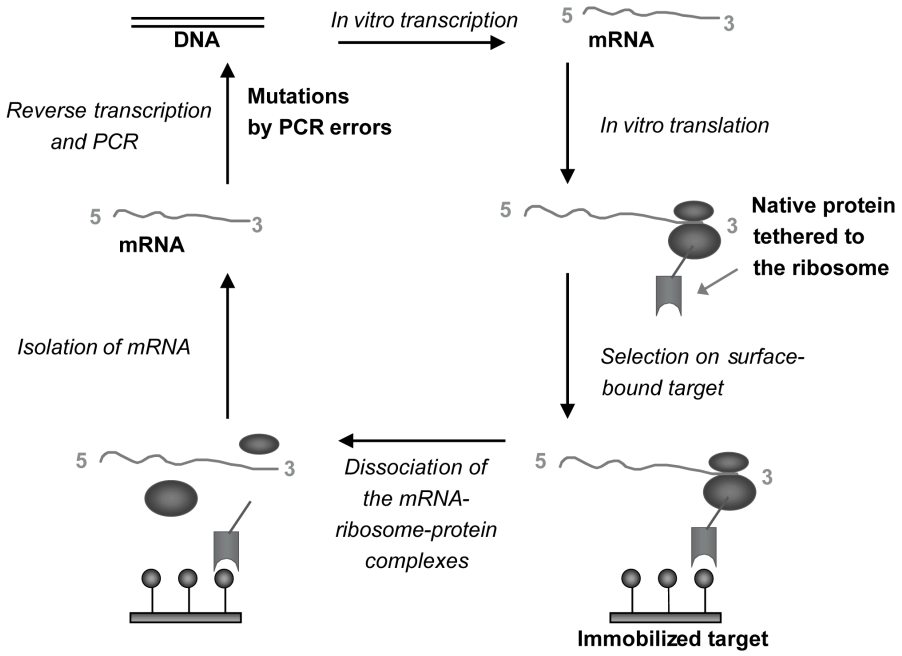


Figure 11. Principle of ribosome display. In the first step the DNA library is amplified by PCR and DNA is transcribed to RNA. RNA is then translated *in vitro* and the desired ribosome complexes are affinity selected by binding to immobilised ligand. The bound ribosome complexes are dissociated and RNA is isolated from the complexes and reverse transcribed to complementary DNA. Amplified DNA is used for the next cycle of selection. Figure reprinted by permission from Macmillan Publishers Ltd: [Nature Biotechnology] Hanes *et al.* (2000).

AIMS OF THE STUDY

The aim of the present study was to apply and develop modified avidin proteins for biotechnology applications. In the first part of the study, the interactions between avidin and azo ligands were studied. In the study, both the structure and function of avidin were modified.

More specifically, the aims were:

I. To study and characterize the interaction between avidin/AVR4 and azo-compounds.

II. To produce and analyze the characteristics of second-generation dual-chain avidins.

III. To change the ligand-binding properties of avidin to generate novel affinity proteins based on the avidin scaffold.

IV. To examine the properties of chimeric avidin as a bioactive molecule covalently immobilized on cellulose acetate film.

3. SUMMARY OF METHODS

The materials and methods are described in more detail in the original publications (I-IV) and the references therein.

3.1 Modification of the avidin gene

In this study, both the structural and functional properties of avidin were changed using protein engineering techniques. The nucleotide sequences of the constructs used were verified by sequencing on an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems) according to the protocols recommended by the manufacturer (ABI PRISM BigDye Terminator Cycle Sequencing Kit v.1.1, Applied Biosystems).

3.1.1 Structural modifications (II)

Circular permutations. To create chimeric dual-chain avidin DNA constructs of circularly permutated avidin cp54 was fused to the circularly permuted genes for streptavidin, avidin-related protein 2 (AVR2) and avidin-related protein 4 (AVR4) (cp65 *SA/AVR2/AVR4*). The cp65 strategy was applied as described in Nordlund *et al*.. Circularly permuted genes were inserted into the pBVboostFG plasmid (Laitinen *et al.* 2005) containing the region encoding the *OmpA* signal sequence, and the plasmid was transformed into chemically competent *E. coli* TOP10 cells (Invitrogen) by the heat-shock method as described in Section 3.2.

3.1.2 Functional modifications (IV)

Construction of a phagemid vector. The avidin gene was first subcloned into the pCR®2.1-TOPO plasmid by TOPO TA-cloning (Invitrogen) and transformed into *E. coli* TOP10 cells. Plasmids were isolated from colonies that were determined to contain inserts based on blue-white screening. Fragments were digested from the pCR®2.1-TOPO -plasmid using the *NheI* and *NotI* restriction enzymes and ligated into the phagemid vector (pBluescript SK+ -derived phagemid, VTT Technical Research Center of Finland, Espoo, Finland). *AVD* cDNA the was subcloned into the phagemid vector as an N-terminal fusion to the C-terminal domain (amino acids 198-406) of the minor phage coat protein III. The avidin phagemid vectors were transformed into chemically competent *E. coli* XL1-Blue cells (Stratagene) by the heat-shock method as described in Section 3.2.

Construction of the Avd L1,2 library. To construct a Avd L1,2 library, amino acids N12, D13, L14, G15, and S16 in the loop between beta strands 1 and 2 of avidin were randomized. The libraries were constructed using a two-step PCR-strategy essentially as previously described (Barbas *et al.* 2001). Fragments were ligated into a phagemid vector (pBluescript SK+ -derived phagemid, Research Center of Finland, Biotechnology, Espoo, Finland), and the resulting ligation product was transformed into electrocompetent cells of the *E. coli* strain XL1-Blue (Stratagene) by electroporation as described in Section 3.2.

Construction of the sbAvd L3,4 library. A DNA library based on steroid-binding avidin (sbAvd), was constructed and ligated into the phagemid to improve the specificity of sbAvd. In the created library (sbAvd library L3,4), amino acids T35, A36, V37, and T38 in the loop between beta-strands 3 and 4 of avidin were randomized. Library construction was performed as described above.

3.2 *Escherichia coli* strains and transformation methods

***Escherichia coli* strains.** The bacterial strain TOP10 [*F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(Str^R) endA1 λ$^-$*] (Invitrogen) was used for plasmid amplification. Strain BL21-AI^{TM} [*F$^-$ ompT hsdS$_B$ (r$_B^-$m$_B^-$) gal dcm araB::T7RNAP-tetA*] (Invitrogen) was used for protein expression. The XL1-Blue strain [*endA1 gyrA96(nal^R) thi-1 recA1 relA1 lac glnV44 F'[::Tn10 proAB$^+$ lacI^q Δ(lacZ)M15] hsdR17(r$_K^-$ m$_K^+$)*] (Stratagene) was used for phage amplification.

Transformations. Plasmids were transformed into bacterial cells by the heat-shock method or by electroporation. Heat-shock was used for standard transformations following the manufacturer's instruction (Invitrogen). Electroporation was used to create avidin gene libraries. Electroporation was preformed on ice with a Gene Pulser (Bio-Rad) following the instructions for the transformation of XL1-Blue cells (Stratagene).

3.3 Selection of steroid-binding proteins (VI)

Avidin display on M13 phage. The functionality of the avidin phages was first tested by competing avidin and avidin(N118M) phages. Amplification of the phages and panning on surfaces coated with BSA-conjugated HABA (HABA-BSA) (Hofstetter *et al.* 2000) was performed essentially as previously described (Barbas *et al.* 2001). Vigorous shaking in hydrochloric acid containing D-biotin (Biochemica, Fluka, 14400) was used for elution. In total, three selection rounds were performed.

Selection of steroid-binding proteins. Steroid-binding avidins were selected from the Avd (L1,2) library by panning against a testosterone-coated surface. The panning procedure was essentially performed as previously described (Barbas *et al.* 2001). Three to four selection cycles were conducted, and acid treatment in the presence of the ligand was used for elution. After each panning round, the results were verified by sequencing, and the number of phage particles was determined by phage titration.

Improvement of the specificity of sbAvd. More specific steroid-binding avidins were selected from sbAvd library L3,4 by panning against a testosterone-coated surface. The panning procedure was essentially performed as previously described (Barbas *et al.* 2001). Three to four selection cycles were conducted, and acid treatment with testosterone was used for elution. After each panning round, results were verified by sequencing, and the number of phage particles was determined by phage titration. Additionally, every panning round was screened for binding proteins using anti-M13 as previously described (Kingsbury, Junghans 1995).

3.4 Expression and purification of recombinant proteins (II, IV)

The pET101-Directional TOPO (Invitrogen) was used as an expression vector. Recombinant proteins were produced in *E. coli* strain BL21-AI (Invitrogen). Small-size cultivations were set up as previously described (Hytönen *et al.* 2004).

Fermentation. The Infors Labfors 3 fermentor was used for protein fermentation. LB medium containing gentamycin and glucose was used for fermentation, supplemented with antifoam agent struktol J647 (Schill & Seilacher, Hamburg, Germany). At the beginning of fermentation, the pO_2 of dissolved oxygen was 0 %, stirring speed 500 rpm and OD_{600} 0.288. Fermentation was performed at 28 °C. The level of oxygen was slowly raised to obtain 20 % dissolved oxygen (pO_2). A feedback loop to maintain the oxygen level at 20 % was set by adjusting the stirring speed between 150 and 1100 rpm. Pumping of the feed solution (50 % glycerol, 2.5 g/l $MgSO_4$, 33 ml/h) was initiated at OD ~1. rotein production was induced at OD ~1.5 and continued for 24 h.

Purification by affinity chromatography. Bacterial pellets were lysed by sonication (Sonics & Materials Vibra Cell™) and DNaseI (New England Bio labs) treatment. Proteins were purified from bacterial lysate by affinity chromatography using 2-iminobiotin (Affiland) (as previously described (Hytönen *et al.* 2004)), biotin (Affiland), or polyhistidine-tag (QIAGEN) columns. For His-tag -purification ÄKTA-

purifier™ HPLC was used. The concentration of purified protein was measured with a NanoDrop™ 3300 spectrometer, and samples were analyzed by SDS-PAGE gels stained with Coomassie Brilliant Blue.

3.5 Summary of structural and functional analyses of proteins (I-IV)

Table 5. Methods used for protein structure and ligand-binding analyses.

Structural analyses	**Instruments/ programs**	**Publication**	**Reference(s)**
Gel-filtration	ÄKTApurifier™ HPLC	**II, IV**	
Thermal stability analysis (SDS-PAGE)		**II**	Bayer, Ehrlich-Rogozinski & Wilchek (1996)
Thermal stability analysis (DSC)	Nano II DSC, capillary VP-DSC	**I, II, IV**	Nordlund *et al.* (2003)
Antibody staining - rabbit α-avd IgG with goat anti-rabbit IgG - mouse anti-pIII IgG with goat anti-mouse IgG		**II, IV**	University of Oulu; Sigma-Aldrich, Biosite, Sweden
Determination of 3D -structure	Beamline X13 (EMBL) XDS, TLS, PROCHECK, WHATIF	**I**	PDB code 1VYO Laskowski *et al.* (1993) Laskowski *et al.* (1993) Vriend (1990)
Modeling	Modeller 9v2, BODIL	**I, II, IV**	Eswar *et al.* (2006) Lehtonen *et al.* (2004)
Docking and MD simulations	Sybyl, Gold 2.2 CHARMM 22, NAMD 2.6, VMD 1.8.6, PyMOL 1.3	**I,II**	Phillips *et al.* (2005) Humphrey, Dalke & Schulten (1996)
Functional analyses	**Instruments /programs**	**Publication**	**Reference(s)**
Microplate analysis	Bio-Rad 680 XR	**IV**	Barbas *et al.* (2001)
Fluorescence-labeled biotin assay	QuantaMaster™	**II**	Hytönen *et al.* (2004)
Molecular recognition force microscopy (MRFM)	Pico SPM I	**IV**	Kamruzzahan et al. (2004) Wildling et al. (2009)
Surface plasmon resonance (SPR)	Biacore X, BiaEvaluation software	**II, IV**	
UV/Visible spectroscopy	PerkinElmer spectrometer	**I**	

3.6 Synthesis of azo ligands (I)

A set of azo dyes were synthesized by coupling diazotized aminobenzoic acid to phenol and naphthol derivatives. During the synthesis of 1,2-dihydroxyazobenzenecarboxylic acid, 1,2,3-dihydroxyazobenzene carboxylic acid and 1,2-dihydroxy-3-methylazobenzenecarboxylic acid, hydroxyl groups were protected by aluminum sulfate before reaction. The resulting precipitates were collected, washed with water and dried *in vacuo*. The product was recrystallized from the ethanol-water mixture and dried. The final products were confirmed using ^{1}H NMR (Bruker Avance DPX 250 FT), ^{13}C NMR (DPX 500) and ESI-MS.

3.7 Generation of bioactive cellulose acetate films (III)

Functionalization of the films. Cellulose acetate (CA) films were first amino-functionalized by dip-coating the films into a sol-gel solution. Next, films were treated with a glutaraldehyde solution. After glutaraldehyde treatment, the films were washed with water, dried and stored at room temperature prior to use. Gas chromatography was used to measure the density of free amino groups on the amino functionalized CA film and amino-reactive groups on the glutaraldehyde functionalized CA film. The amino and amino-reactive groups on the surface of functionalized films were reacted with benzaldehyde and aniline analytes, respectively, to form imine bonds. Scanning electron microscopy (SEM) images of prepared CA films were generated using Zeiss ULTRA plus field emission microscope.

Avidin proteins (chimeric avidin, chicken avidin, rhizavidin, streptavidin) were added onto the films by incubating the films in the appropriate protein solution. Coated films were analyzed using tritium-labeled biotin as described below. The chimeric avidin-functionalized films were stored either at RT or at +4^{o}C for three months, and the activity of these films was analyzed monthly.

Determination of biotin-binding activity. The biotin-binding capacity of the surface-immobilized avidin layer was measured using tritium-labeled biotin (D-[8,9-^{3}H]-biotin, Amersham) with an LKB Wallac 1217 RackBeta liquid scintillation counter (Turku, Finland). Non-specific biotin binding was taken into account by measuring the bound [^{3}H]-biotin after blocking the films with an excess of D-biotin. These results were subtracted from the results without D-biotin blocking to calculate the specific biotin-binding capacity. All biotin-binding capacity results were calculated as bound biotin per area of the film [mol/cm^2].

4. SUMMARY OF RESULTS

4.1 High-resolution crystal structure of avidin (I)

Avidin contains a classical beta-barrel structure with eight antiparallel beta-strands (Figure 12). The X-ray structure of avidin has been previously described (Livnah *et al.* 1993b; Pugliese *et al.* 1993; Livnah *et al.* 1993a; Huberman *et al.* 2001; Nardone *et al.* 1998; Pazy *et al.* 2003; Pazy *et al.* 2002), but in this study, the structure of avidin was determined with high resolution (1.48 Å). In the determined structure, both the open and closed conformations of the loop between β-strands 3 and 4 were observed.

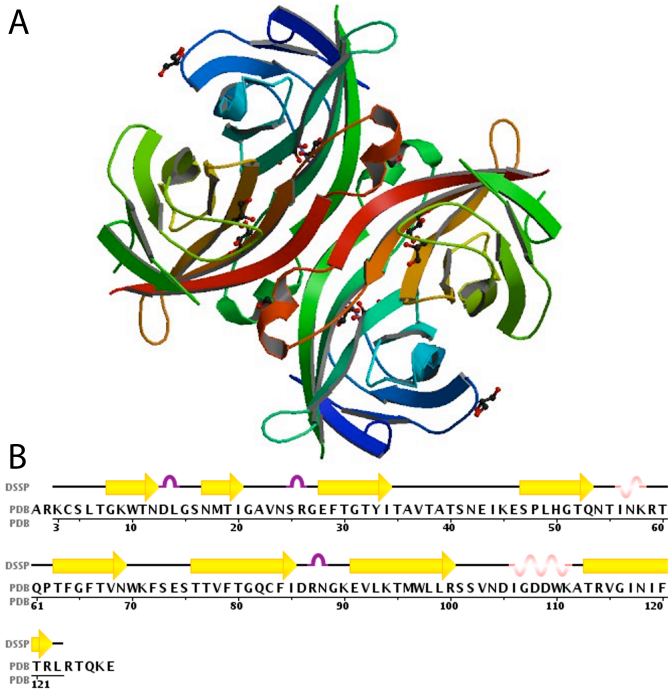


Figure 12. The structure and sequence of avidin (PDB code 1VYO). (A) The determined structure of tetrameric avidin. Two glycerol molecules in the ligand-binding site are shown as ball and stick models. (B) Primary and secondary structures of avidin.

4.2 Structural modification of avidin (II)

To overcome challenges limiting the further use of the dual-chain avidin, chimeric dual-chain avidins, tandem fusions of avidin and a homologous protein (streptavidin (SA), avidin-related protein 2 (AVR2), or avidin-related protein 4 (AVR4)), were generated. All of these proteins were selected as chimeric fusion partners because they resemble avidin in their fold and 3D-structure (Weber *et al.* 1987; Eisenberg-Domovich *et al.* 2004; Hytonen *et al.* 2005a). Fusion proteins were modeled based on the previously determined 3D -structures of dcAVD (PDB 2C4I), SA (PDB 1MK5), and AVRs [AVR2 (PDB 1WBI), AVR4 (PDB 1Y53)]. Molecular dynamics (MD) were performed, and loop dynamics were analyzed by root mean square fluctuation (RMSF) (Figure 13). In the analysis, we observed that the linkers connecting the original termini of the domains were highly mobile. Additionally, in the tandem fusion of dcAVD and SA, there were unique mobile regions in the loop between strands β7 and β8 (residues 164-169) and in the loop between strands β4 and β5 (residues 243-249), as compared to other tandem fusions.

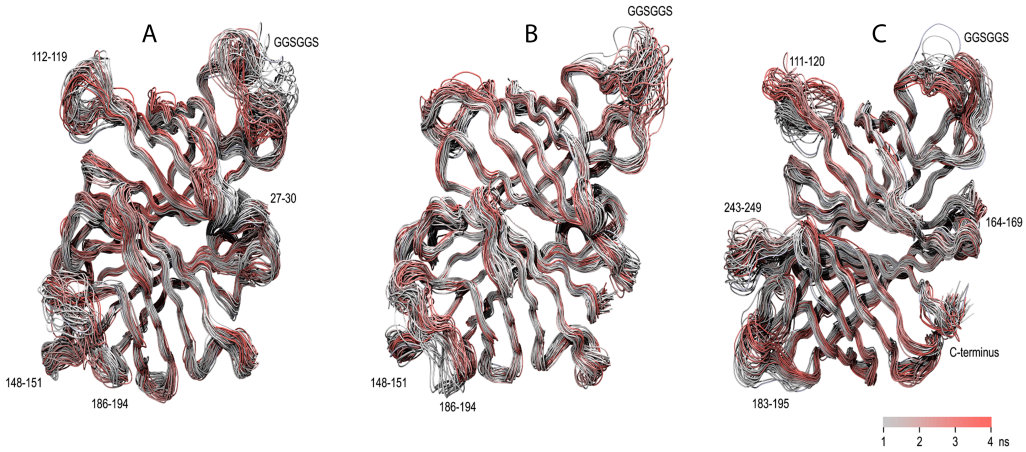


Figure 13. Molecular dynamic simulation of local dynamics in the chimeric dcAVD fusions. The dynamics of the structure are illustrated by plotting 50 superimposed structural snapshots along the 4-ns simulation (A: dcAVD/AVR2; B: dcAVD/AVR4; C: dcAVD/SA). The loops showing a high amount of structural fluctuation (RMSF > 1.5Å) are indicated by numbers referring to the amino acid sequence.

In the case of the tandem fusion of avidin and AVR4, an increase in protein production (5 mg of pure protein per liter of production medium) and improved thermal stability (Table 6) were observed, as compared to the original dcAVD. Additionally,

PCR amplification of the hybrid gene was more efficient, thus enabling more convenient and straightforward modification of the dual-chain avidin. When the ligand-binding properties of dcAVD/AVR4 were studied, biphasic biotin dissociation was detected.

Table 6. Transition temperatures for subunit dissociation (T_r) and thermal unfolding (T_m) determined as SDS-PAGE analysis and DSC.

	SDS-PAGE		**DSC**	
	T_r (°C)	**$T_r^{(+BTN)}$ (°C)**	**T_m (°C)**	**$T_m^{(+BTN)}$ (°C)**
AVD	60[a]	95[a]	84[b]	117[b]
AVR4(C112S)	90	95	110[b]	127[b]
dcAVD	40	75[c]	80[a]	116[a]
dcAVD/AVR4	70	85	*86*/91[d]	*108*/112[d]

[a]Value from(Hytönen *et al.* 2004); [b](Hytönen *et al.* 2004); [c](Nordlund *et al.* 2004);

[d](two-peak analysis was applied to the data, and the minor peak is shown in *italics*).

$^{+BTN}$ = with biotin.

4.3 Functional modification of avidin (IV)

The chicken-derived protein avidin is known for its extremely high affinity for the water-soluble vitamin H, D-biotin. To alter or change this extremely tight interaction, the avidin gene was subjected to oligonucleotide-directed mutagenesis, and avidin gene libraries were created. To select avidin mutants with novel binding properties from the created gene libraries, a phage display method was used. Avidin was displayed monovalently (3+3) on the surface of the M13 phage as a single fusion with the C-terminal region of the minor coat protein p3.

To generate avidin gene libraries, the loops adjacent to the ligand-binding site were selected for randomization (Livnah *et al.* 1993b). In the libraries, four to nine amino acids were randomized (Table 7) using both NNN and NNY degenerate codons. The NNN codon covers all 20 naturally occurring amino acids, whereas the use of the NNY codon limits the sequence space of the library by covering only 14 amino acids, thus eliminating all of the stop codons.

Table 7. Avidin gene libraries generated and the randomized amino acids for each.

Library	Size $\times 10^6$	Theoretical size $\times 10^6$	Randomized residues	Degenerate codon
Avd L1,2	0.1	3.2	**N12**,D13,L14,G15,**S16**	NNN
sbAvd-1 L3,4	1.4	0.38	**T35**,A36,V37,**T38**	NNY
Avd L5,6*	2.1	0.16	**F72**,**S73**,E75,**S75**	NNN
Avd L1,2+L5,6*	5.9	20661	**N12**,D13,L14,G15,**S16** + **F72**,**S73**,E75,**S75**	NNY

The library size is calculated from the transformation efficiency. Randomized amino acids that form hydrogen bonds with biotin are shown in bold. *Unpublished results.

4.3.1 Steroid-binding avidin capture

In the first phage display panning rounds, a combination of acid and biotin was used for elution, but in the final panning round, testosterone was used instead of biotin to elute phage. During selection, the Avd L1,2 loop library phages were introduced onto testosterone-coated surfaces. The quality of the phage genomes carrying the mutated cDNA was evaluated by DNA sequencing after every panning round. The sequence logo depicted in figure 14 shows the most frequent amino acid residues in the randomized area. A clear enrichment of sequences in the randomized loop area was observed, indicating the success of the selection conditions. An Avd variant named sbAvd-1, carrying the sequence N12, R13, M14, N15, H16 was selected for further analysis.

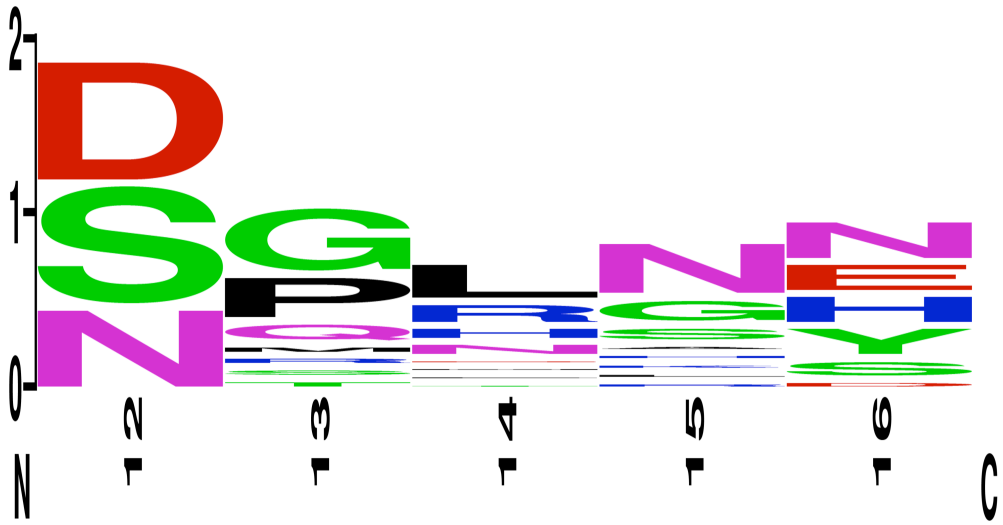


Figure 14. Sequence logo from panning of Avd L1,2 library (http://weblogo.berkeley.edu/ 20.8.2011) describes the most frequent amino acids present in the randomized area. Logo is combination of 11 sequences from 3rd panning round. The most conserved was amino acid number 12; asparagines (N), aspartic acid (D), and serine (S) were most frequently present in this position. Wt avidin has amino acid sequence NDLGS in the corresponding site.

To improve the specificity of sbAvd-1, the sbAvd-1 L3,4 library was screened by phage display panning on a testosterone-coated surface. Several different panning procedures were used with varying preincubation, washing, and elution conditions. Generally, washes were adjusted according to the number of output colonies, and a combination of acid and testosterone was used for elution. The quality of the phage genomes carrying the mutated cDNA was evaluated by DNA sequencing. Additionally, phage clones were evaluated by microplate analysis using BSA-testosterone as a target ligand and using anti-M13 antibody to determine the amount of bound phage. The sbAvd-1 variant that showed the highest binding activity in the microplate assay had the sequence A35, T36, V37, N38 and was selected for further analyses in addition to several other clones (Table 8).

Table 8. Sequencing results from one round of sbAvd-1 L3,4 panning with testosterone (unpublished data). SbAvd-1 mutants captured from this panning experiment carrying the sequences $V^{35}Y^{36}D^{37}Y^{38}$, $V^{35}N^{36}P^{37}V^{38}$, $V^{35}V^{36}T^{37}H^{38}$, and $L^{35}L^{36}R^{37}G^{38}$ were produced and analyzed by microplate analysis. These mutants are marked in bold in the table.

Panning round	Sample	Elution	T^{35}	A^{36}	V^{37}	T^{38}
1	1	TES	Y	P	T	R
1	2	TES	V	F	T	P
1	3	TES	N	T	L	A
1	4	TES	I	Y	T	V
1	5	TES	A	S	T	R
1	6	TES	I	R	C	Y
1	7	TES	P	F	S	L
1	8	TES	R	N	V	H
3	1	TES	N	L	I	C
3	2	TES	L	P	C	F
3	3	TES	C	L	P	N
3	4	TES	T	L	L	D
3	5	TES	F	F	H	A
3	6	TES	V	I	I	Y
3	7	TES	L	T	C	T
3	8	TES	F	I	T	T
3	9	TES	V	I	I	L
3	10	TES	S	Y	W	L
3	11	TES	R	F	N	A
3	12	TES	I	G	D	A
3	13	TES	G	G	V	R
3	14	TES	S	N	S	L
4	1	TES	**V**	**Y**	**D**	**Y**
4	2	TES	V	I	S	H
4	3	TES	T	P	D	P
4	4	TES	F	N	P	C
4	5	TES	C	R	A	R
4	6	TES	**V**	**N**	**P**	**V**
4	1	DHEAS	**V**	**V**	**G**	**R**
4	2	DHEAS	H	F	L	C
4	3	DHEAS	G	D	G	D
4	4	DHEAS	V	V	T	H
4	5	DHEAS	C	D	R	L
4	6	DHEAS	**L**	**L**	**R**	**G**
4	7	DHEAS	D	C	H	I
4	8	DHEAS	S	Y	Y	V
4	9	DHEAS	L	R	P	H

TES = testosterone, DHEAS = dehydroepiandrosterone sulfate

4.3.2 Characteristics of steroid-binding avidins

Proteins captured from the phage display selections were produced in soluble tetrameric form (Table 9.) in the *E. coli* strain BL21-AI an using N-terminal OmpA bacterial secretion signal from *Bordetella avium* (Hytönen *et al.* 2004). Proteins were purified by Ni-NTA affinity chromatography, which yielded ~2 mg/L of pure protein. The produced proteins were screened using microplate assays, and the most promising candidates (sbAvd-1, sbAvd-2) were subjected to further analyses.

Table 9. Properties of steroid-binding avidins as analyzed by gel-filtration, SPR, DLS, and MRFS.

Protein	**Gel-filtration** (kDa)	**Ligand**	**SPR** k_a (1/Ms)	k_d (1/s)	K_D (μM)	**DSC** T_m (˚C)	ΔT_m (˚C)	**MRFS** binding events
wtAvd	53[a]	-				85.5		
		BNT	n.d.	n.d.	n.d.	123.2	37.7	
		TES		no binding		86.2	0.7	
sbAvd-1	51	-				80.6		183
		BTN	420000	0.0006	0.0014	83.2	2.6	
		TES	1000	0.0095	9.0	81.5	0.9	67
sbAvd-2	53	-				82.5		215
		BTN	1000	0.0007	0.7	83.0	0.5	
		TES	813	0.0085	11	83.1	0.6	21

[a]Result from (Hytönen *et al.* 2004).

Based on the differential scanning calorimetric (DSC) assays, the high thermal stability of wt avidin (temperature transition midpoint (T_m) = 85.5 °C) was preserved in the steroid-binding mutants; these proteins showed T_m values similar to that of wt avidin (sbAvd-1 T_m value of 80.6 °C and sbAvd-2 T_m value of 82.5 °C, Table 9).

Molecular recognition force spectroscopy (MRFS) (Hinterdorfer, Dufrene 2006) was used to study the interaction between sbAvds and testosterone on the single-molecule level. In these experiments, tip-tethered testosterone was found to form a complex with the sbAvds, and the most frequent unbinding force (Baumgartner, Hinterdorfer & Schindler 2000) was found to be similar for both of the testosterone-sbAvd complexes: 40 pN at a constant pulling velocity of 600 nm/s. From 1,000 recorded force-distance cycles, sbAvd-1 showed 183 detected interactions, and sbAvd-2 showed 215 events (Table 9). Additionally, when the binding sites of the

proteins were preloaded with free testosterone, the number of binding events dropped to 67 from 183 in the case of sbAvd-1. In the case of sbAvd-2, the number of interactions dropped from 215 to 21 after the addition of free testosterone. These results indicate the specificity of the testosterone binding.

The kinetic constants of testosterone and biotin binding to the steroid-binding proteins were determined using surface plasmon resonance (SPR). Purified sbAvd-1 and sbAvd-2 bound to a testosterone-BSA-coated sensor chip with comparable affinities (Table 9), whereas wt avidin showed no binding to the testosterone-coated surface. Furthermore, testosterone binding was inhibited by varying testosterone concentrations (0.75-50 μM). In the case of sbAvd-1, 50% inhibition was achieved between 5 and 10 μM testosterone, which is consistent with the determined testosterone-coated surface-binding affinity. Interestingly, the sbAvd-2 measurements suggested a considerably higher binding affinity for free testosterone than that measured for surface-immobilized testosterone because 50% inhibition was already achieved at a testosterone concentration of 750 nM.

Biotin binding by the steroid-binding proteins was analyzed using a biotin-coated sensor chip. A comparison of sbAvd-1 and sbAvd-2 revealed that the biotin-binding affinity decreased almost 500-fold due to the modification of the loop between *β*-strands 3 and 4.

The binding specificity of the proteins was analyzed by microplate assays, in which a set of different small molecules (progesterone, PRO; hydrocortisone, HYD; testosterone, TES; cholic acid, CHO) were used as target molecules (Figure 15). Wt avidin on a biotin-coated surface was used as a positive control along with sbAvd proteins on a BSA-conjugated testosterone-coated surface (TES-B). The negative controls PBS, BSA, and HSA did not show any response. sbAvd-1 and sbAvd-2 showed clear binding to testosterone and progesterone, and binding to the surface-immobilized ligands was inhibited by preincubating the proteins with 10 μM D-biotin (Figure 15 (+BTN)). Free biotin significantly inhibited the binding of sbAvd-1 to steroids, indicating a notable affinity of the protein for biotin. However, in the case of sbAvd-2, the binding of testosterone and progesterone was not affected by biotin, showing a clear decrease in biotin-binding affinity.

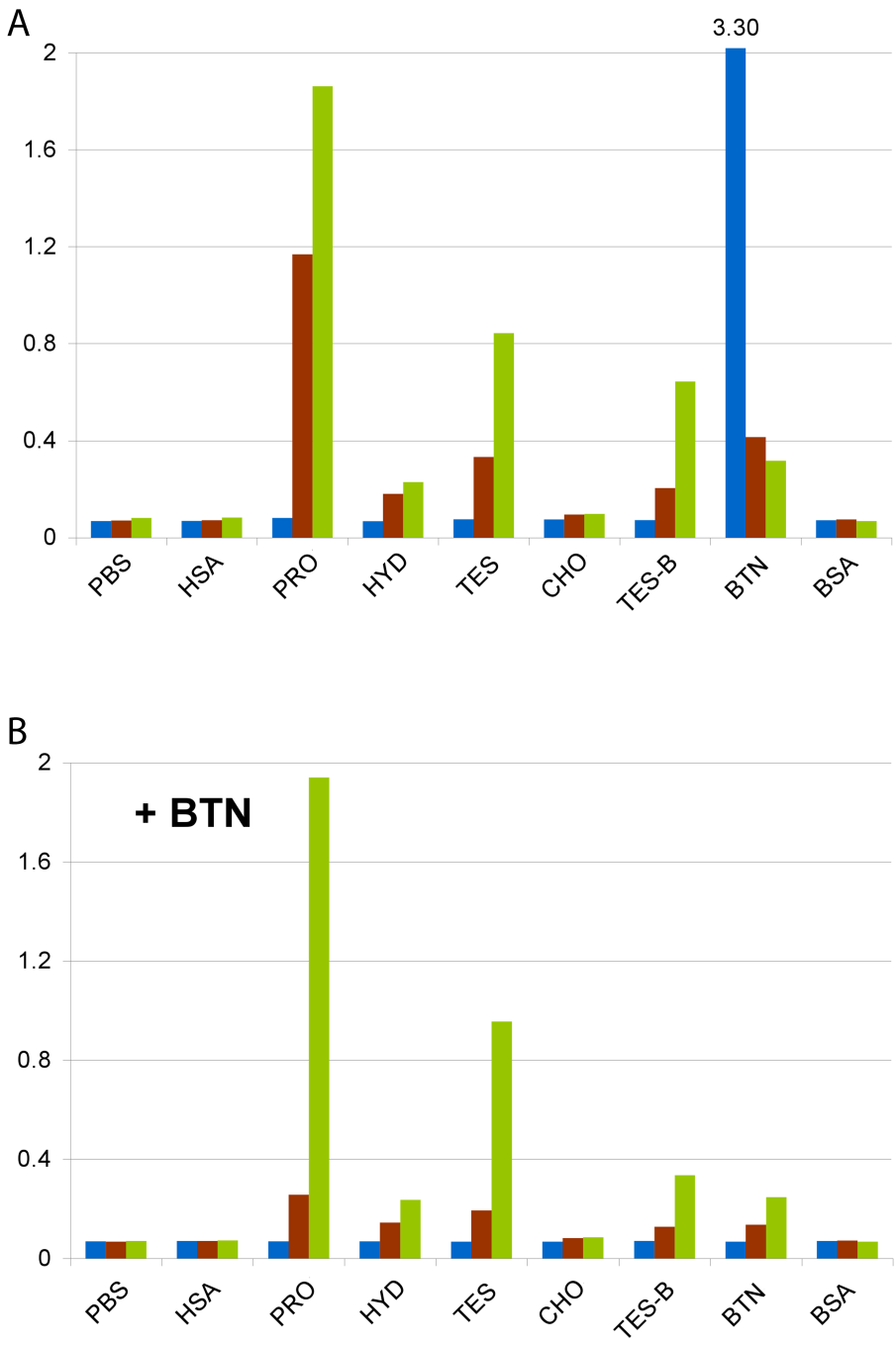


Figure 15. Determination of ligand-binding specificity of sbAvd-1 and sbAvd-2 proteins by microplate analysis. The binding of wt avidin (blue bars), sbAvd-1 (red bars), and sbAvd-2 (green bars) to a set of different small ligands was detected using polyclonal anti-avidin antibody as a probe (A). Also the effect of free biotin (10 µM) to ligand-binding was analyzed (B). Abbreviations used in the figure: PBS, Phosphate buffered saline; HSA, human serum albumin; PRO, HSA-conjugated progesterone; HYD, HSA-conjugated hydrocortisone; TES, HSA-conjugated testosterone; CHO, HSA-conjugated cholic acid; TES-B, BSA-conjugated testosterone; BTN, BSA-conjugated D-biotin; BSA, bovine serum albumin

The specificity of steroid binding was also analyzed using SPR by competing biotin with the following steroids: testosterone, dehydroepiandrosterone sulfate (DHEAS), androstenedione, estradiol, and dihydrotestosterone (DHT). As expected, sbAvd-1 and sbAvd-2 behaved differently in the experiment. sbAvd-1 was found to be cross-reactive with androgens highly similar to testosterone (DHEAS and androstenedione) and had a detectable affinity for biotin. Interestingly, DHT showed less efficient inhibition than testosterone, suggesting a lower binding affinity for sbAvd-1 for this steroid. The binding of sbAvd-2 to testosterone was found to be most efficiently inhibited by testosterone and DHEAS (Figure 16), whereas the inhibition caused by biotin was clearly weaker than in the case of sbAvd-1. This result is consistent with the microplate analysis results and with the binding kinetic constants determined using SPR.

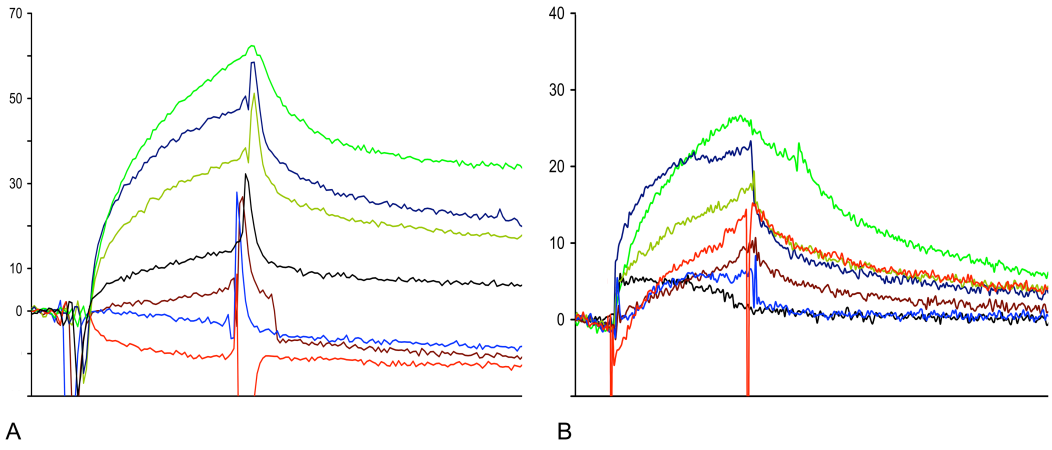
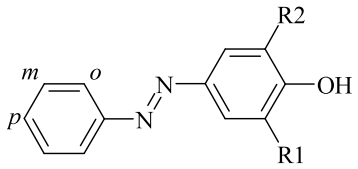


Figure 16. Inhibition analysis of sbAvd-1 and sbAvd-2 proteins analyzed by the SPR method. The binding of the sbAvd-1 and the sbAvd-2 to a CM5 sensor chip functionalized with testosterone-BSA was measured in the presence of 50 µM inhibitors. (A) The binding of the sbAvd-1 protein was totally inhibited by DHEAS, androstenedione, and biotin. (B) The binding of the sbAvd-2 protein was totally inhibited by DHEAS or testosterone, but not with biotin. This indicates markedly decreased affinity of the sbAvd-2 protein towards biotin. Samples: Protein sample, green; protein with estradiol, dark blue; protein with DHT, olive; protein with testosterone, black; protein with DHEAS, blue; protein with androstenedione, brown; protein with biotin, red

4.4 Binding of azo-ligands to avidin and AVR4 (I)

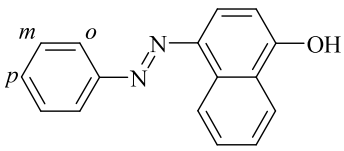
Aside from biotin, avidin also binds dyes and peptides that share no significant structural similarity to biotin (Green 1975, Green 1970). The binding of HABA-derived azo dyes to avidin and AVR4 was studied using DSC, UV/visible spectroscopy, and molecular modeling. For this purpose a set of 15 azo dye compounds (Table 10) were synthesized and characterized.

Table 10. Functional groups of the synthesized azo compounds.



Compound	-COO$^-$ position	R1	R2
1b	*meta*	H	H
1c	*para*	H	H
2b	*meta*	H	NO_2
2c	*para*	H	NO_2
4a	*ortho*	CH_3	CH_3
4b	*meta*	CH_3	CH_3
4c	*para*	CH_3	CH_3
5a	*ortho*	H	OH
5b	*meta*	H	OH
5c	*para*	H	OH
6b	*meta*	OH	CH_3
6c	*para*	OH	CH_3

p=para, m=meta, o=ortho



Compound	-COO$^-$ position
3a	*ortho*
3b	*meta*
3c	*para*

By DSC, HABA and five other azo dyes were found to significantly increase the stability of avidin (ΔT_m<10˚C). In the case of AVR4, the stabilization effect was not as strong as for avidin. However, for both proteins, the highest stabilization effect was provided by the 1-naphthol derivative compound 3a (ΔT_m=15.9˚C for avidin, ΔT_m=7.0˚C for AVR4).

According to UV/visible spectroscopy measurements, the addition of avidin or AVR4 caused spectral changes. After the addition of avidin, the largest spectral changes were detected for compounds 3a, 3c, 4a, and 5a. However, in the case of AVR4, the spectra of compounds 1c, 3a, 4a, 5a, and 5c were the most significantly changed among the compounds measured.

4.5 Generation and analysis of bioactive films (III)

The functionalization of cellulose acetate (CA) films was performed in two steps (Figure 17). Bioactivity was generated on the films by immobilization of avidin proteins (chimeric avidin, chicken avidin, rhizavidin, streptavidin) onto the CA films.

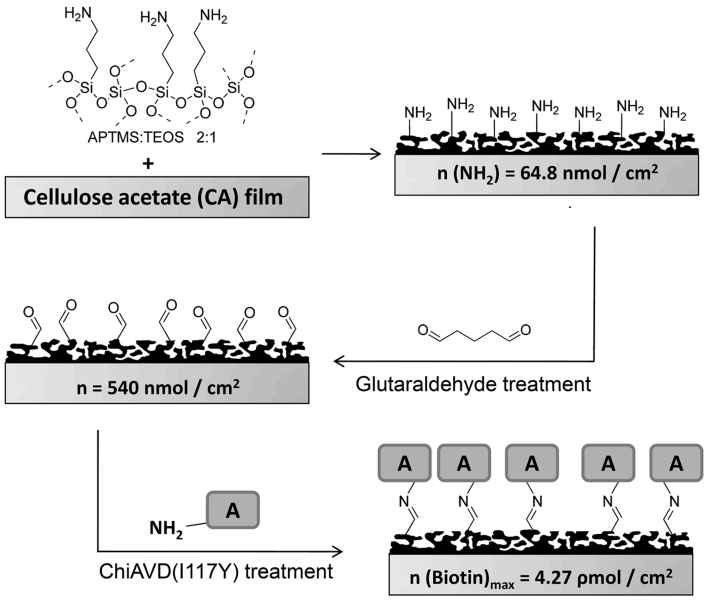


Figure 17. A schematic representation of functionalization and protein immobilization on the cellulose acetate surface.

The specific biotin-binding capacity of the films was determined by comparing the total biotin-binding capacity with the results obtained after blocking the surfaces

with D-biotin before the [^{3}H]-biotin assay. Importantly, no significant differences were observed in the binding capacities of control films (R1-R3) (Figure 18).

For the chimeric avidin-coated films, the binding capacity was analyzed monthly throughout the three-month study period (Figure 18). At first, both non-functionalized film (A1) and glutaraldehyde-activated film (A3) bound approximately the same amount of [^{3}H]-biotin, indicating that the same amount of chimeric avidin was immobilized onto both films. The amino-functionalized film (A2) was found to have approximately 50% less immobilized chimeric avidin compared to the A1 and A3 films. The A3 film was found to be the most stable; covalently bound avidin remained fully active throughout the study in both storage conditions tested.

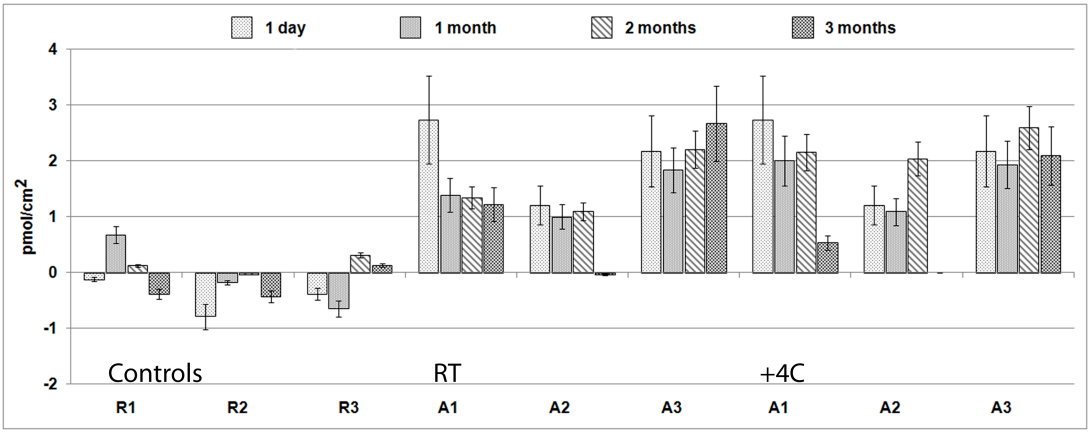


Figure 18. The specific biotin-binding capacity of the CA films measured with [^{3}H]-biotin assay. The control films (R1=non-functionalized film, R2=amino-functionalized film, R3=glutaraldehyde activated film) contained no avidin and as expected showed no biotin-binding activity. In the study the most stable film was glutaraldehyde activated CA films with chimeric avidin coating (A3). (The amino-functionalized film with chimeric avidin coating (A2) stored at +4°C for 3 months was not analyzed.)

In the analysis of the films coated with chicken avidin, streptavidin and rhizavidin, the films coated with chicken avidin showed similar, high biotin-binding capacities to those detected in the case of chimeric avidin. Surprisingly, films coated with streptavidin and rhizavidin had low biotin-binding capacities.

5. DISCUSSION

5.1 The structure of avidin provides a promising scaffold for protein engineering

Chicken avidin has been widely used in life science research applications (Laitinen *et al.* 2007). Avidin provides an attractive and robust scaffold for the development of novel receptors because avidin has many properties required of a novel affinity scaffold, such as high chemical and thermal stability, and an oligomeric nature enabling signal amplification. The ligand-binding site of avidin resembles the CDRs of antibodies and is optimized for the binding of small molecules. Additionally, the structure of avidin is well characterized, and numerous engineered forms of avidin have been described (Laitinen *et al.* 2006).

To expand our knowledge of the avidin structure and ligand binding, the high-resolution crystal structure of avidin was solved in the first part of this study (I). The new structure revealed interesting details about the loop between beta-strands 3 and 4, as in the previously determined structure of avidin (Livnah *et al.* 1993a), the loop is not visible. The loop between beta-strands 3 and 4 of wt avidin has high mobility, whereas in the other protein structure studied (AVR4), the flexibility of the loop is constrained by a salt bridge (Eisenberg-Domovich *et al.* 2005). For that reason, the mouth of the binding-pocket of AVR4 is wider than that of avidin. In this study, the spectral properties of azo compounds were studied by UV/visible spectroscopy in the absence and presence of the avidin and AVR4 proteins. Avidin is known to significantly affect the absorption spectrum of HABA (Green 1965) because HABA forms conformational tautomers called hydrazone tautomers when bound to avidin (Livnah *et al.* 1993a). From the performed experiments, it can be concluded that avidin and AVR4 have different binding strategies; avidin prefers *ortho* derivatives for binding, whereas AVR4 prefers *para* derivatives of azo compounds.

5.1.1 Structural modification of avidin: dcAVD/AVR4

A good example of the complicated genetic engineering possible with the robust structure of avidin is dual-chain avidin (dcAVD) (Nordlund *et al.* 2004, Hytonen *et al.* 2005b). DcAVD is generated avidin form, in which the two pairs of the binding sites can be independently manipulated. Also forms in which all four binding sites (single-chain avidin) can be independently manipulated, have been created (Nordlund *et al.* 2005). However, the PCR amplification of the dcAVD-encoding sequence has been challenging because primers recognize complementary sequences from both subunits, which would produce several different PCR products. This same issue also limits targeted mutagenesis of dcAVD.

In the second part of the study (II), the usability of dcAVD was improved by creating chimeric dual-chain avidin consisting of fusions of wt avidin to streptavidin, AVR2 or AVR4. The chimeric tandem fusion of circularly permutated wt avidin and circularly permuted AVR4 showed enhanced protein expression and thermal stability when compared to the original dcAVD. Additionally, the PCR amplification was more straightforward using the chimeric dual chain fusion gene. Interestingly, the binding analyses of dcAVD/AVR4 protein showed that the molecule exhibited heterogenous biotin-binding, and it was concluded that the exchange of the cp65-subunit of dcAVD with a circularly permuted AVR4 subunit lowered biotin-binding affinity of the AVR4-derived cp65-subunit.

DcAVD molecules offer a valuable base for development of bi-functional avidin tools for various applications such as bioseparation. This kind of study also provides novel information of the unfolding mechanisms of avidin proteins.

5.1.2 Steroid-binding avidin prefers steroids to biotin as ligands

Changing the binding properties of a high-affinity molecule without altering the structure is a challenging task. The natural ligand of avidin, biotin, is a water-soluble vitamin that is composed of a ureido ring fused with a tetrahydrothiophene ring (Figure 19). A valeric acid substituent is attached to one of the carbon atoms of the tetrahydrothiophene ring. Testosterone is a steroid hormone derived from cholesterol (Figure 19). Steroid hormones are hydrophobic molecules containing only a few functional groups. Therefore, the ligand-binding site for a steroid-binding mole-

cule should be a deep hydrophobic cavity, where shape complementary is generated through aromatic side chains with appropriately placed hydrogen bond donors and acceptors (Schlehuber, Skerra 2005).

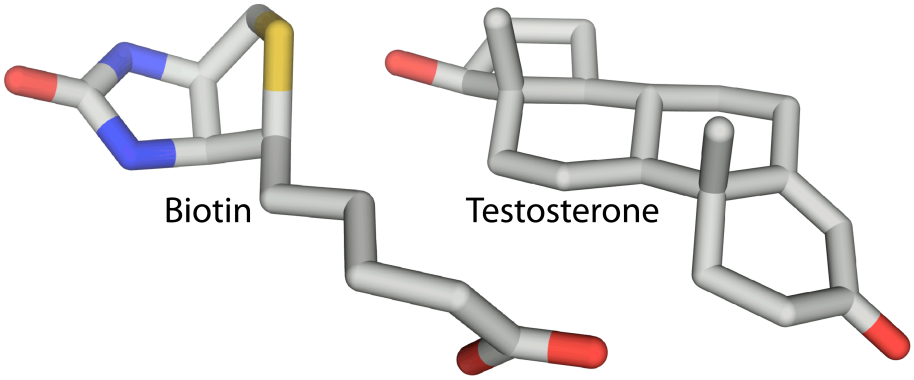


Figure 19. The structure of D-biotin and testosterone.

Although steroid hormones are, in many ways, highly problematic small molecules, several antibodies have been generated to recognize them with high affinity. The majority of commercially available immunodiagnostic kits for steroid hormones are based on polyclonal rabbit antibodies because of the lack of appropriate monoclonal antibodies (Hemminki 1998). Wilson *et al.* studied the binding of progesterone by the DB3 anti-progesterone antibody (Arevalo *et al.* 1993). This antibody has a nanomolar affinity for progesterone. However, it is also highly cross-reactive with progesterone-like steroids. Testosterone is recognized by the anti-testosterone antibody 3-C_4F_5 with subnanomolar affinity, and its cross-reactivity was decreased by stepwise optimization of the CDRs (Hemminki *et al.* 1998a, Hemminki *et al.* 1998b). Antibodies that bind the hydrophilic steroid digoxigenin, frequently used in diagnostic applications, have been generated, and recently, a bispecific antibody with digoxigenin specificity has been introduced (Metz *et al.* 2011). Digoxigenin can be coupled to low molecular weight compounds such as fluorophores, chelating agents, nucleic acids, lipids, nanoparticles, and proteins (Kessler 1991).

As an alternative to antibodies, molecules recognizing digoxigenin have been generated from the lipocalin fold (Figure 20) (Schlehuber, Beste & Skerra 2000a). Lipocalins have beta-barrel structures similar to avidin, and these proteins belong to the same calycin superfamily (Flower 1996). Therefore, it was likely that a steroid-binding avidin variant could be engineered.

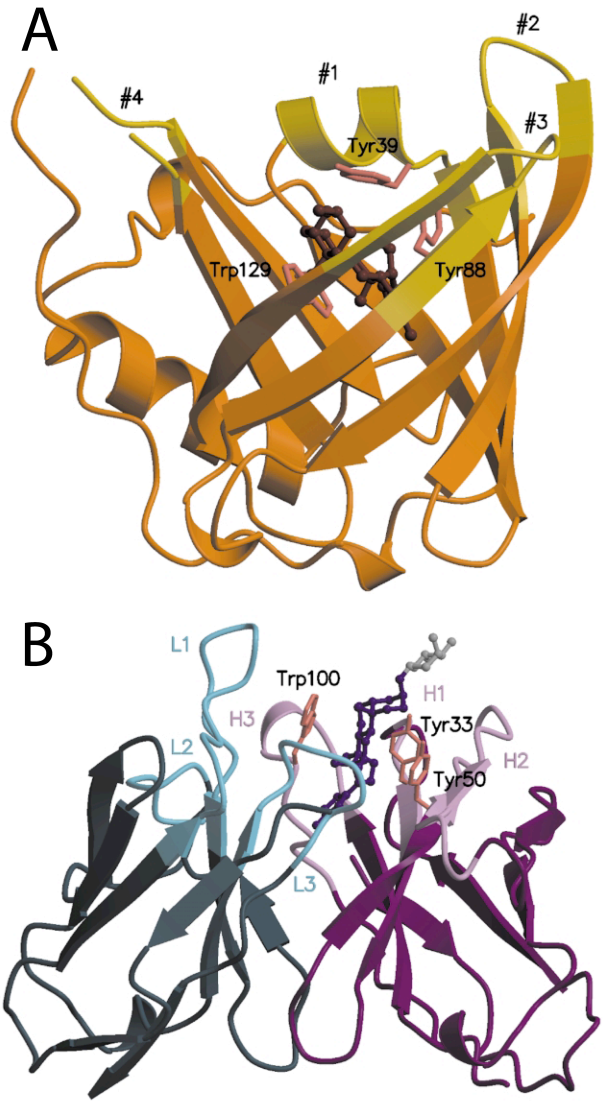


Figure 20. Comparison of the anti-digoxigenin antibody with the engineered lipocalin DigA16. (A) The structure of DigA16 with ligand. The hypervariable loop segments are colored yellow and the three aromatic side-chains of Tyr39, Tyr88, and Trp129 are shown in pink. (B) Fv fragment of the antibody in complex with digoxin (violet, at the centre; only one digitoxose sugar attached to C-3 is visible and colored gray). The V_L and V_H domains are colored green and violet, respectively, while the CDRs are depicted in corresponding lighter colors. Figure is reprinted from Korndorfer, Schlehuber & Skerra (2003) with permission from Elsevier.

In the present study (VI), the loop regions of avidin were subjected to targeted random mutagenesis to change avidin into a steroid-binding molecule. At first, five residues in the loop between β-strands 1 and 2 of avidin were totally randomized to generate a population of diverse avidin genes. From the phage display selections

(Smith 1985) the enrichment of avidin variants binding to testosterone was observed and an avidin mutant (sbAvd-1) that carried the sequence N12, R13, M14, N15, H16, was further analyzed and found to bind free testosterone with a micromolar affinity. However, the biotin-binding affinity of the protein was still high enough to inhibit sbAvd-1 from binding to testosterone-BSA.

To decrease biotin binding and to improve steroid specificity sbAvd-1 was further engineered. From the phage display selections, an sbAvd-1-derived protein, named sbAvd-2, with a mutated loop between β-strands 3 and 4 was captured. This mutant showed similar or slightly decreased binding affinity for immobilized testosterone, whereas the biotin-binding affinity was clearly decreased. We also noticed that sbAvd-2 bound more tightly to free testosterone than did sbAvd-1. In many cases, the determined affinities for immobilized ligands are lower than those determined for free ligands because of the different conditions used in the binding assays and due to ligand accessibility to the binding site. Significantly, sbAvd-2 was found to prefer steroids as a ligand to biotin, and the thermostability of the avidin was preserved in the developed molecules.

However, there are still factors that may limit the use of the developed molecule, for instance, in diagnostic applications. Cross-reactivity of the molecules with steroids other than testosterone would be highly problematic because the number of closely related steroid structures in serum samples is high. In addition, the micromolar affinity of steroid-binding avidins would be too low for diagnostic applications. The *in vivo* concentrations of testosterone are low, down to the picomolar range. Additionally, the steroid-binding proteins were rather insoluble and were prone to aggregation, which would limit the large-scale production of these proteins. If the developed affinity proteins were used for therapeutic use, the problem of potential immunogenicity would be an issue to consider because the avidin protein is of foreign origin and will likely cause an immune response if no precaution is taken.

5.2 Bioactive films

In the third part of the study (III), a new method for the biofunctionalization of cellulose acetate films using a simple two-step protocol was developed. In the biofunctionalization process, chimeric avidin was utilized for its stable structure in

harsh conditions and also for its efficient and scalable expression in *E. coli*. The determined specific biotin-binding capacity of glutaraldehyde-functionalized films coated with chimeric avidin was 2.17 pmol/cm^2. This value is clearly lower than that of a fully active avidin monolayer (6.6 pmol/cm^2). If the obtained values are converted to molecular density, there would be 0.33 biotin binding sites per surface area of an avidin tetramer. This suggests an average coverage of approximately 33% or less of the surface coated with functional chimeric avidins.

The developed material offers a valuable platform for the development of inexpensive *in vitro* diagnostics and also supports biosensing applications that are required to operate under demanding conditions. However, the advantages to the use of chimeric avidin instead of avidin should be carefully evaluated because there are already available diagnostic applications based on the avidin scaffold.

5.3 Future directions

5.3.1 Avidin gene libraries and selection strategies

During construction of the gene libraries, two bottlenecks limited the library size: 1) the ligation of the insert into the digested phagemid vector and 2) the transformation. Traditional ligation was rather ineffective, and a considerable amount of DNA was required for successful ligation. Transformation was performed by electroporation, which is a more powerful method than the heat-shock method. However, in this method, a limited amount of DNA can be added into the cells, thus limiting the size of the library. In the study, libraries up to 6×10^6 individuals were created, but increasing the efficiency of the ligation, for example with a Gateway-based cloning system, would result in larger libraries (Hiltunen S. *et al*., unpublished results).

To be able to cover the entire sequence space and, at the same time, randomize several amino acids, a binary code or another variation-decreasing codon method (Table 4) should be used in the library design. A gene library with restricted diversity provides knowledge about the capacity of different amino acid residues for mediating molecular recognition. As already mentioned in Section 2.1.1.1, the importance of serine, glycine and tyrosine in antigen recognition by antibodies has been discovered using binary code (Birtalan *et al.* 2008).

The phage display method was used as a selection system in study IV. At first, avidin was displayed on the phage surface. The successful display of avidin on the phage surface was expected because avidin can be efficiently expressed in a soluble form in *E. coli* (Hytönen *et al.* 2004), and the close relative of avidin, streptavidin, has been displayed on bacterial phages (Avrantinis *et al.* 2002; Sidhu, Weiss & Wells 2000). Two different avidin display modes were evaluated, in which the first developed mode produced only p3 fusions and the second produced both p3 fusions and avidin without phage protein. However, these methods did not include an amber stop codon-based system, which would enable production of the tetrameric protein and, thus, increase the display of avidin on the phage surface. Another challenge in the avidin phage display method was the culture medium, which contained biotin. Preliminary results show that the addition of free wt avidin in the phage amplification step increases the amount of active phages.

5.3.2 Steroid-binding avidins

Engineered steroid-binding avidins (IV) could be used, for example, in bioseparation applications, but as discussed in section 5.1.2, there are limiting factors for the use of the mutants in diagnostic applications. Encouragingly, however, the automated panning performed with multiple ligands conjugated to human serum albumin (Antti Tullila, VTT Technical Research Center of Finland) showed successful results (Hiltunen S. *et al.*, unpublished data). These results suggest that it will be possible to further develop affinity molecules from the avidin scaffold.

Finally, techniques developed in this study could be combined. The most successful dcAVD-based molecule, dcAVD/AVR4, could be combined with other genetically engineered avidins with modified or completely novel binding properties. By combining sbAvd with a dcAVD avidin platform, it would be possible to develop avidin-based receptors with alternative binding affinities or with multiple ligand specificities to be used in *in vitro* diagnostics or in nanotechnology

6. SUMMARY AND CONCLUSION

In this thesis, chicken avidin, a protein extensively utilized in the life science applications, was modified to produce novel molecules for avidin-biotin technology. First, the high-resolution 3D-structure of avidin was resolved, and binding to 15 azo molecules was analyzed. From the structure and ligand binding analysis of avidin and AVR4, the importance of the loop between β-strands 3 and 4 was observed. This loop was shown to seal the ligand-binding site in the closed conformation, thus acting as an important coordinator in ligand binding by avidin.

Avidin was structurally modified by fusing circularly permuted avidin and circularly permuted homologous protein in tandem to overcome the challenges associated with the original dual-chain avidin. The chimeric dual-chain avidin, a tandem fusion of avidin and avidin-related protein 4 (AVR4), showed increased protein production levels and increased thermal stability compared to the original dual-chain avidin. PCR amplification of the hybrid gene was also more efficient, thus enabling more convenient and straightforward modification of the dual-chain avidin. The generated dual-chain molecule offers a valuable base for developing bifunctional avidin tools. Additionally, this strategy could be helpful when generating hetero-oligomers from other oligomeric proteins with high structural similarity.

The ligand-binding function of avidin was changed in this study; steroid binding-avidin was created by combinatorial protein engineering techniques. The selection of a sequence-randomized avidin library led to the isolation of a steroid-binding avidin mutant (sbAvd-1) showing micromolar affinity for testosterone. Furthermore, the sbAvd-1 protein was further engineered to increase the specificity of the molecule, and a steroid-binding protein with significantly decreased biotin-binding affinity was obtained. Importantly, the properties that make avidin an attractive molecule for engineering, such as high thermal and chemical stability, were preserved in the engineered avidins.

A novel and simple two-step protocol was developed for the biofunctionalization of cellulose acetate films. Highly thermostable chimeric avidin, which is a geneti-

cally engineered version of the high-affinity biotin-binding protein avidin, was immobilized on the films. The activity of the cellulose acetate with immobilized chimeric avidin was retained completely for three months, even when stored at room temperature. Avidin-coated cellulose acetate films could be used, for example, in the development of inexpensive and sensitive diagnostic tools to be used in personalized medicine platforms.

In conclusion, this study presents the 3D -structure of avidin and analyses of the azo ligand-binding properties of the protein. Two different avidin modifications were performed; as a structural modification, a chimeric tandem fusion of dual-chain avidin was created, and as functional modifications, steroid-binding avidins were obtained. Additionally, this study describes the development of biofunctionalized films that utilize the molecular recognition properties of the avidin molecule.

ACKNOWLEDGEMENTS

I wish to express my great expression of gratitude to Professor Markku Kulomaa who has been supervisor in this study. He has provided many interesting opportunities to broaden my knowledge and to grow as a scientist. Additionally, I want to thank Markku, for always having first-class working environment and facilities.

I wish to thank to my other supervisor Docent Vesa Hytönen for professional guidance through this journey. His enthusiastic attitude towards science and his endless thirst for knowledge has been very inspiring. I want to thank Soili Hiltunen for the great contribution to my research. I am confident that this research will go on within the most proficient hands. Many thanks to Sampo Kukkurainen; you have done great work with modeling and simulations. Warm thanks to my colleague and friend Dr. Juha Määttä. We have shared many good laughs together, and I wish that there are still many to come. I want to thank Tiia Koho for sharing the important things (also other than science) with me. It has been a great counterbalance to work. I wish to express my gratitude to Ulla Kiiskinen for the excellent technical support. I would like to thank also other MBT and PD group members, the new and the former ones: late Docent Henri Nordlund, Barbara Niederhauser, Jenni Leppiniemi, Dr. Jenita Pärssinen, Outi Väätäinen, Laura Kananen, Dr. Jarkko Valjakka, Dr. Satu Helppolainen, Dr. Anssi Mähönen, Tiina Tissari, Dr. Olli Laitinen (Vactech) and, Suvi Varjonen (VTT). It has been pleasant to work with you!

I started this research at VTT Technical Research Center of Finland in the guidance of my thesis committee member Professor Kristiina Takkinen in 2005. During six-month laboratory visit at VTT, I learnt phage display-methodology, which has been the corner stone of my research. I would like to thank Kristiina for kind and professional guidance into the phage world. Furthermore, I would like to thank VTT lab technicians, especially Pirkko Veijola-Bailey and Armi Boman, for their valuable tips and instructions. I am delighted that the collaboration between MBT/PD and VTT has continued. I thank Antti Tullila and Dr. Tarja Nevanen for sharing ideas and for the successful collaboration.

During these years I have had pleasure to supervise two Masters Theses and guide summer students. Thank you, Soili Hiltunen, Johanna Koivisto, Hong Chang, Markus Ojanen, Elina Ojala, Johanna Ketolainen, and Hanna-Mari Seppälä for your valuable contribution to my research. Additionally, I wish to thank people from other IBT groups for broadening my knowledge about other things than avidin: Thank you Janette Hinkka for mentoring high school students with me. Thank you Dr. Katri Köninki for valuable discussions and sparkling peer-support.

I would like to thank reviewers Professor Kari Airenne and Dr. Petri Saviranta for their valuable comments and constructive criticism during writing process. Additionally, I wish to express my sincere gratitude to all my collaborators and co-authors not already mentioned: Professor Peter Hinterdorfer, Professor Osmo Hormi, Professor Kari Rissanen, Docent Juhani Huuskonen, Dr. Andreas Ebner, Dr. Teemu Ihalainen, Dr. Einari Niskanen, Dr. Susanna Repo, Dr. Thomas Nyholm, Dr. Jarkko Heikkinen, Martina Rangl, Jarno Hörhä, and Aulikki Lehmus.

This research has been supported by Tampere Graduate Program in Biomedicine and Biotehcnology (TGPBB), European Micro and Nano Technology support program (project number FFG 421695), and the Academy of Finland (project numbers 115976 and 121236). I thank also Pirkanmaa Hospital District for financial support.

I am grateful to my dear friends, to my relatives and to my family-in-law for their support and caring. I wish to express a great expression of gratitude to my father for excellent support during my studies. You have always been there for me. I want to thank also my sister, Ardalan, and little Noah for love and support. I wish that the future brings us to the same country. Kiitokset tuesta rakkaalle Mumpsille, olet ihana!

My deepest gratitude goes to my dear husband Tuomo and to my lovely daughter Pihla. I am so lucky to have You in my life, I love You!

This PhD thesis is dedicated to Pihla, you are the sunshine of my life!

REFERENCES

Aggarwal, S. 2010, "What's fueling the biotech engine-2009-2010", *Nature biotechnology,* vol. 28, no. 11, pp. 1165-1171.

Arevalo, J.H., Stura, E.A., Taussig, M.J. & Wilson, I.A. 1993, "Three-dimensional structure of an anti-steroid Fab' and progesterone-Fab' complex", *Journal of Molecular Biology,* vol. 231, no. 1, pp. 103-118.

Avrantinis, S.K., Stafford, R.L., Tian, X. & Weiss, G.A. 2002, "Dissecting the streptavidin-biotin interaction by phage-displayed shotgun scanning", *Chembiochem,* vol. 3, no. 12, pp. 1229-134.

Barbas, C.F., III, Burton, D.R., Scott, J.K. & Silverman, G.J. (eds) 2001, *Phage Display, A Laboratory manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Barbas, C.F.,3rd, Kang, A.S., Lerner, R.A. & Benkovic, S.J. 1991, "Assembly of combinatorial antibody libraries on phage surfaces: the gene III site", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 88, no. 18, pp. 7978-7982.

Baumgartner, W., Hinterdorfer, P. & Schindler, H. 2000, "Data analysis of interaction forces measured with the atomic force microscope", *Ultramicroscopy,* vol. 82, no. 1-4, pp. 85-95.

Bayer, E.A., Ehrlich-Rogozinski, S. & Wilchek, M. 1996, "Sodium dodecyl sulfate-polyacrylamide gel electrophoretic method for assessing the quaternary state and comparative thermostability of avidin and streptavidin", *Electrophoresis,* vol. 17370, pp. 1319-1324.

Beck, A., Haeuw, J.F., Wurch, T., Goetsch, L., Bailly, C. & Corvaia, N. 2010a, "The next generation of antibody-drug conjugates comes of age", *Discovery medicine,* vol. 10, no. 53, pp. 329-339.

Beck, A., Wurch, T., Bailly, C. & Corvaia, N. 2010b, "Strategies and challenges for the next generation of therapeutic antibodies", *Nature reviews.Immunology,* vol. 10, no. 5, pp. 345-352.

Beste, G., Schmidt, F.S., Stibora, T. & Skerra, A. 1999, "Small antibody-like proteins with prescribed ligand specificities derived from the lipocalin fold", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 96, no. 5, pp. 1898-1903.

Binz, H.K., Amstutz, P., Kohl, A., Stumpp, M.T., Briand, C., Forrer, P., Grutter, M.G. & Pluckthun, A. 2004, "High-affinity binders selected from designed ankyrin repeat protein libraries", *Nature biotechnology,* vol. 22, no. 5, pp. 575-582.

Binz, H.K., Amstutz, P. & Pluckthun, A. 2005, "Engineering novel binding proteins from nonimmunoglobulin domains", *Nature biotechnology,* vol. 23, no. 10, pp. 1257-1268.

Birtalan, S., Zhang, Y., Fellouse, F.A., Shao, L., Schaefer, G. & Sidhu, S.S. 2008, "The intrinsic contributions of tyrosine, serine, glycine and arginine to the af-

finity and specificity of antibodies", *Journal of Molecular Biology,* vol. 377, no. 5, pp. 1518-1528.
Bostrom, J., Yu, S.F., Kan, D., Appleton, B.A., Lee, C.V., Billeci, K., Man, W., Peale, F., Ross, S., Wiesmann, C. & Fuh, G. 2009, "Variants of the antibody herceptin that interact with HER2 and VEGF at the antigen binding site", *Science (New York, N.Y.),* vol. 323, no. 5921, pp. 1610-1614.
Boulianne, G.L., Hozumi, N. & Shulman, M.J. 1984, "Production of functional chimaeric mouse/human antibody", *Nature,* vol. 312, no. 5995, pp. 643-646.
Cadwell, R.C. & Joyce, G.F. 1992, "Randomization of genes by PCR mutagenesis", *PCR methods and applications,* vol. 2, no. 1, pp. 28-33.
Chothia, C., Lesk, A.M., Tramontano, A., Levitt, M., Smith-Gill, S.J., Air, G., Sheriff, S., Padlan, E.A., Davies, D. & Tulip, W.R. 1989, "Conformations of immunoglobulin hypervariable regions", *Nature,* vol. 342, no. 6252, pp. 877-883.
Clackson, T. & Wells, J.A. 1994, "In vitro selection from protein and peptide libraries", *Trends in biotechnology,* vol. 12, no. 5, pp. 173-184.
Clark, L.A., Ganesan, S., Papp, S. & van Vlijmen, H.W. 2006, "Trends in antibody sequence changes during the somatic hypermutation process", *Journal of immunology (Baltimore, Md.: 1950),* vol. 177, no. 1, pp. 333-340.
Colas, P., Cohen, B., Jessen, T., Grishina, I., McCoy, J. & Brent, R. 1996, "Genetic selection of peptide aptamers that recognize and inhibit cyclin-dependent kinase 2", *Nature,* vol. 380, no. 6574, pp. 548-550.
Cwirla, S.E., Peters, E.A., Barrett, R.W. & Dower, W.J. 1990, "Peptides on phage: a vast library of peptides for identifying ligands", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 87, no. 16, pp. 6378-6382.
Daugherty, P.S. 2007, "Protein engineering with bacterial display", *Current opinion in structural biology,* vol. 17, no. 4, pp. 474-480.
DeLange, R.J. & Huang, T.-. 1971, "Egg white avidin. III. Sequence of the 78-residue middle cyanogen bromide peptide. Complete amino acid sequence of the protein subunit", *The Journal of Biological Chemistry,* vol. 2469, pp. 698-709.
DeLano, W.L. 2002, "Unraveling hot spots in binding interfaces: progress and challenges", *Curr Opin Struct Biol,* vol. 12, no. 1, pp. 14-20.
Deng, L.W., Malik, P. & Perham, R.N. 1999, "Interaction of the globular domains of pIII protein of filamentous bacteriophage fd with the F-pilus of Escherichia coli", *Virology,* vol. 253, no. 2, pp. 271-277.
Dennis, M.S. & Lazarus, R.A. 1994, "Kunitz domain inhibitors of tissue factor-factor VIIa. I. Potent inhibitors selected from libraries by phage display", *The Journal of biological chemistry,* vol. 269, no. 35, pp. 22129-22136.
Desmyter, A., Decanniere, K., Muyldermans, S. & Wyns, L. 2001, "Antigen specificity and high affinity binding provided by one single loop of a camel single-domain antibody", *The Journal of biological chemistry,* vol. 276, no. 28, pp. 26285-26290.
Duenas, M. & Borrebaeck, C.A. 1994, "Clonal selection and amplification of phage displayed antibodies by linking antigen recognition and phage replication", *Bio/technology (Nature Publishing Company),* vol. 12, no. 10, pp. 999-1002.
Eakin, R.E., Snell, E.E. & Williams, R.J. 1941, "The concentration and assay of avidin, the injury-producing protein in raw egg-white", *Journal of biological chemistry,* vol. 140, pp. 535-543.

Ehrenstein, M.R. & Notley, C.A. 2010, "The importance of natural IgM: scavenger, protector and regulator", *Nature reviews.Immunology,* vol. 10, no. 11, pp. 778-786.

Eisenberg-Domovich, Y., Hytonen, V.P., Wilchek, M., Bayer, E.A., Kulomaa, M.S. & Livnah, O. 2005, "High-resolution crystal structure of an avidin-related protein: insight into high-affinity biotin binding and protein stability", *Acta crystallographica.Section D, Biological crystallography,* vol. 61, no. Pt 5, pp. 528-538.

Eswar, N., Webb, B., Marti-Renom, M.A., Madhusudhan, M.S., Eramian, D., Shen, M.Y., Pieper, U. & Sali, A. 2006, "Comparative protein structure modeling using Modeller", *Current protocols in bioinformatics,* vol. Chapter 5, pp. Unit 5.6.

Fellouse, F.A., Li, B., Compaan, D.M., Peden, A.A., Hymowitz, S.G. & Sidhu, S.S. 2005, "Molecular recognition by a binary code", *Journal of Molecular Biology,* vol. 348, no. 5, pp. 1153-1162.

Fellouse, F.A., Wiesmann, C. & Sidhu, S.S. 2004, "Synthetic antibodies from a four-amino-acid code: a dominant role for tyrosine in antigen recognition", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 101, no. 34, pp. 12467-12472.

Flower, D.R. 1996, "The lipocalin protein family: structure and function", *The Biochemical journal,* vol. 318 (Pt 1), no. Pt 1, pp. 1-14.

Flower, D.R. 1993, "Structural relationship of streptavidin to the calycin protein superfamily", *FEBS Lett,* vol. 333, no. 1-2, pp. 99-102.

Flower, D.R., North, A.C. & Sansom, C.E. 2000, "The lipocalin protein family: structural and sequence overview", *Biochimica et biophysica acta,* vol. 1482, no. 1-2, pp. 9-24.

Gebauer, M. & Skerra, A. 2009, "Engineered protein scaffolds as next-generation antibody therapeutics", *Current opinion in chemical biology,* vol. 13, no. 3, pp. 245-255.

Geisberger, R., Lamers, M. & Achatz, G. 2006, "The riddle of the dual expression of IgM and IgD", *Immunology,* vol. 118, no. 4, pp. 429-437.

Getts, D.R., Getts, M.T., McCarthy, D.P., Chastain, E.M. & Miller, S.D. 2010, "Have we overestimated the benefit of human(ized) antibodies?", *mAbs,* vol. 2, no. 6, pp. 682-694.

Giver, L. & Arnold, F.H. 1998, "Combinatorial protein design by in vitro recombination", *Current opinion in chemical biology,* vol. 2, no. 3, pp. 335-338.

Gotz, M., Hess, S., Beste, G., Skerra, A. & Michel-Beyerle, M.E. 2002, "Ultrafast electron transfer in the complex between fluorescein and a cognate engineered lipocalin protein, a so-called anticalin", *Biochemistry,* vol. 41, no. 12, pp. 4156-4164.

Gramatikoff, K., Georgiev, O. & Schaffner, W. 1994, "Direct interaction rescue, a novel filamentous phage technique to study protein-protein interactions", *Nucleic acids research,* vol. 22, no. 25, pp. 5761-5762.

Green, N.M. 1975, "Avidin", *Advances in protein chemistry,* vol. 295, pp. 85-133.

Green, N.M. 1970, "Spectrophotometric determination of avidin and streptavidin", *Methods Enzymol.,* vol. 18, pp. 418-424.

Green, N.M. 1965, "A Spectrophotometric Assay for Avidin and Biotin Based on Binding of Dyes by Avidin", *Biochemical journal,* vol. 94, pp. 23C-24C.

Greenwood, J., Willis, A.E. & Perham, R.N. 1991, "Multiple display of foreign peptides on a filamentous bacteriophage. Peptides from Plasmodium falciparum

circumsporozoite protein as antigens", *Journal of molecular biology,* vol. 220, no. 4, pp. 821-827.
Gronwall, C., Snelders, E., Palm, A.J., Eriksson, F., Herne, N. & Stahl, S. 2008, "Generation of Affibody ligands binding interleukin-2 receptor alpha/CD25", *Biotechnology and applied biochemistry,* vol. 50, no. Pt 2, pp. 97-112.
Gronwall, C. & Stahl, S. 2009, "Engineered affinity proteins--generation and applications", *Journal of biotechnology,* vol. 140, no. 3-4, pp. 254-269.
Hamann, P.R., Hinman, L.M., Hollander, I., Beyer, C.F., Lindh, D., Holcomb, R., Hallett, W., Tsou, H.R., Upeslacis, J., Shochat, D., Mountain, A., Flowers, D.A. & Bernstein, I. 2002, "Gemtuzumab ozogamicin, a potent and selective anti-CD33 antibody-calicheamicin conjugate for treatment of acute myeloid leukemia", *Bioconjugate chemistry,* vol. 13, no. 1, pp. 47-58.
Hanes, J. & Pluckthun, A. 1997, "In vitro selection and evolution of functional proteins by using ribosome display", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 94, no. 10, pp. 4937-442.
Hanes, J., Schaffitzel, C., Knappik, A. & Pluckthun, A. 2000, "Picomolar affinity antibodies from a fully synthetic naive library selected and evolved by ribosome display", *Nature biotechnology,* vol. 18, no. 12, pp. 1287-1292.
Harding, F.A., Stickler, M.M., Razo, J. & DuBridge, R.B. 2010, "The immunogenicity of humanized and fully human antibodies: residual immunogenicity resides in the CDR regions", *mAbs,* vol. 2, no. 3, pp. 256-265.
Hecht, M.H., Das, A., Go, A., Bradley, L.H. & Wei, Y. 2004, "De novo proteins from designed combinatorial libraries", *Protein science,* vol. 13, no. 7, pp. 1711-1723.
Hemminki, A. 1998, *Development of recombinant antibodies for diagnostic applications by protein engineering*, Technical Research Center of Finland, Espoo.
Hemminki, A., Niemi, S., Hautoniemi, L., Soderlund, H. & Takkinen, K. 1998a, "Fine tuning of an anti-testosterone antibody binding site by stepwise optimisation of the CDRs", *Immunotechnology,* vol. 4, no. 1, pp. 59-69.
Hemminki, A., Niemi, S., Hoffren, A.M., Hakalahti, L., Soderlund, H. & Takkinen, K. 1998b, "Specificity improvement of a recombinant anti-testosterone Fab fragment by CDRIII mutagenesis and phage display selection", *Protein engineering,* vol. 11, no. 4, pp. 311-319.
Henning, P., Magnusson, M.K., Gunneriusson, E., Hong, S.S., Boulanger, P., Nygren, P.A. & Lindholm, L. 2002, "Genetic modification of adenovirus 5 tropism by a novel class of ligands based on a three-helix bundle scaffold derived from staphylococcal protein A", *Human gene therapy,* vol. 13, no. 12, pp. 1427-1439.
Hey, T., Fiedler, E., Rudolph, R. & Fiedler, M. 2005, "Artificial, non-antibody binding proteins for pharmaceutical and industrial applications", *Trends in biotechnology,* vol. 23, no. 10, pp. 514-522.
Hinterdorfer, P. & Dufrene, Y.F. 2006, "Detection and localization of single molecular recognition events using atomic force microscopy", *Nature methods,* vol. 3, no. 5, pp. 347-355.
Hofstetter, H., Morpurgo, M., Hofstetter, O., Bayer, E.A. & Wilchek, M. 2000, "A labeling, detection, and purification system based on 4-hydroxyazobenzene-2-carboxylic acid: an extension of the avidin-biotin system", *Analytical Biochemistry,* vol. 284, no. 2, pp. 354-366.
Hogbom, M., Eklund, M., Nygren, P.A. & Nordlund, P. 2003, "Structural basis for recognition by an in vitro evolved affibody", *Proceedings of the National*

Academy of Sciences of the United States of America, vol. 100, no. 6, pp. 3191-3196.
Hoogenboom, H.R. & Chames, P. 2000, "Natural and designer binding sites made by phage display technology", *Immunology today,* vol. 21, no. 8, pp. 371-378.
Hoogenboom, H.R., de Bruine, A.P., Hufton, S.E., Hoet, R.M., Arends, J.W. & Roovers, R.C. 1998, "Antibody phage display technology and its applications", *Immunotechnology,* vol. 4, no. 1, pp. 1-20.
Hosse, R.J., Rothe, A. & Power, B.E. 2006, "A new generation of protein display scaffolds for molecular recognition", *Protein science,* vol. 15, no. 1, pp. 14-27.
Huberman, T., Eisenberg-Domovich, Y., Gitlin, G., Kulik, T., Bayer, E.A., Wilchek, M. & Livnah, O. 2001, "Chicken avidin exhibits pseudo-catalytic properties. Biochemical, structural, and electrostatic consequences", *Journal of biological chemistry,* vol. 276, no. 34, pp. 32031-3209.
Hudson, P.J. & Souriau, C. 2003, "Engineered antibodies", *Nature medicine,* vol. 9, no. 1, pp. 129-134.
Humphrey, W., Dalke, A. & Schulten, K. 1996, "VMD: visual molecular dynamics", *Journal of molecular graphics,* vol. 14, no. 1, pp. 33-8, 27-8.
Huse, W.D., Sastry, L., Iverson, S.A., Kang, A.S., Alting-Mees, M., Burton, D.R., Benkovic, S.J. & Lerner, R.A. 1989, "Generation of a large combinatorial library of the immunoglobulin repertoire in phage lambda", *Science,* vol. 246, no. 4935, pp. 1275-1281.
Hytönen, V.P., Laitinen, O.H., Airenne, T.T., Kidron, H., Meltola, N.J., Porkka, E., Hörhä, J., Paldanius, T., Määttä, J.A., Nordlund, H.R., Johnson, M.S., Salminen, T.A., Airenne, K.J., Ylä-Herttuala, S. & Kulomaa, M.S. 2004, "Efficient production of active chicken avidin using a bacterial signal peptide in Escherichia coli", *Biochemical journal,* vol. 384, no. 2, pp. 385-390.
Hytonen, V.P., Maatta, J.A., Kidron, H., Halling, K.K., Horha, J., Kulomaa, T., Nyholm, T.K., Johnson, M.S., Salminen, T.A., Kulomaa, M.S. & Airenne, T.T. 2005a, "Avidin related protein 2 shows unique structural and functional features among the avidin protein family", *BMC biotechnology,* vol. 5, pp. 28.
Hytonen, V.P., Nordlund, H.R., Horha, J., Nyholm, T.K., Hyre, D.E., Kulomaa, T., Porkka, E.J., Marttila, A.T., Stayton, P.S., Laitinen, O.H. & Kulomaa, M.S. 2005b, "Dual-affinity avidin molecules", *Proteins,* vol. 61, no. 3, pp. 597-607.
Hytönen, V.P., Nyholm, T.K., Pentikainen, O.T., Vaarno, J., Porkka, E.J., Nordlund, H.R., Johnson, M.S., Slotte, J.P., Laitinen, O.H. & Kulomaa, M.S. 2004, "Chicken Avidin-related Protein 4/5 Shows Superior Thermal Stability when Compared with Avidin while Retaining High Affinity to Biotin", *Journal of biological chemistry,* vol. 279, no. 10, pp. 9337-9343.
Iannolo, G., Minenkova, O., Petruzzelli, R. & Cesareni, G. 1995, "Modifying filamentous phage capsid: limits in the size of the major capsid protein", *Journal of molecular biology,* vol. 248, no. 4, pp. 835-844.
Junutula, J.R., Flagella, K.M., Graham, R.A., Parsons, K.L., Ha, E., Raab, H., Bhakta, S., Nguyen, T., Dugger, D.L., Li, G., Mai, E., Lewis Phillips, G.D., Hiraragi, H., Fuji, R.N., Tibbitts, J., Vandlen, R., Spencer, S.D., Scheller, R.H., Polakis, P. & Sliwkowski, M.X. 2010, "Engineered thio-trastuzumab-DM1 conjugate with an improved therapeutic index to target human epidermal growth factor receptor 2-positive breast cancer", *Clinical cancer research,* vol. 16, no. 19, pp. 4769-4778.
Kamruzzahan, A.S., Kienberger, F., Stroh, C.M., Berg, J., Huss, R., Ebner, A., Zhu, R., Rankl, C., Gruber, H.J. & Hinterdorfer, P. 2004, "Imaging morphological

details and pathological differences of red blood cells using tapping-mode AFM", *Biological chemistry,* vol. 385, no. 10, pp. 955-960.
Kamtekar, S., Schiffer, J.M., Xiong, H., Babik, J.M. & Hecht, M.H. 1993, "Protein design by binary patterning of polar and nonpolar amino acids", *Science,* vol. 262, no. 5140, pp. 1680-1685.
Kang, A.S., Barbas, C.F., Janda, K.D., Benkovic, S.J. & Lerner, R.A. 1991, "Linkage of recognition and replication functions by assembling combinatorial antibody Fab libraries along phage surfaces", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 88, no. 10, pp. 4363-4366.
Keefe, A.D. & Szostak, J.W. 2001, "Functional proteins from a random-sequence library", *Nature,* vol. 410, no. 6829, pp. 715-718.
Kessler, C. 1991, "The digoxigenin:anti-digoxigenin (DIG) technology--a survey on the concept and realization of a novel bioanalytical indicator system", *Molecular and cellular probes,* vol. 5, no. 3, pp. 161-205.
Kingsbury, G.A. & Junghans, R.P. 1995, "Screening of phage display immunoglobulin libraries by anti-M13 ELISA and whole phage PCR", *Nucleic acids research,* vol. 23, no. 13, pp. 2563-2564.
Knappik, A., Ge, L., Honegger, A., Pack, P., Fischer, M., Wellnhofer, G., Hoess, A., Wolle, J., Pluckthun, A. & Virnekas, B. 2000, "Fully synthetic human combinatorial antibody libraries (HuCAL) based on modular consensus frameworks and CDRs randomized with trinucleotides", *Journal of molecular biology,* vol. 296, no. 1, pp. 57-86.
Kohler, G. & Milstein, C. 1975, "Continuous cultures of fused cells secreting antibody of predefined specificity", *Nature,* vol. 256, no. 5517, pp. 495-497.
Korndorfer, I.P., Schlehuber, S. & Skerra, A. 2003, "Structural mechanism of specific ligand recognition by a lipocalin tailored for the complexation of digoxigenin", *Journal of molecular biology,* vol. 330, no. 2, pp. 385-396.
Krebber, C., Spada, S., Desplancq, D., Krebber, A., Ge, L. & Pluckthun, A. 1997, "Selectively-infective phage (SIP): a mechanistic dissection of a novel in vivo selection for protein-ligand interactions", *Journal of molecular biology,* vol. 268, no. 3, pp. 607-618.
Krebber, C., Spada, S., Desplancq, D. & Pluckthun, A. 1995, "Co-selection of cognate antibody-antigen pairs by selectively-infective phages", *FEBS letters,* vol. 377, no. 2, pp. 227-231.
Krebs, J.E., Goldstein, E.S. & Kilpatrick, S.T. 2011, *Lewin's Genes X*, 10th edn, Jones and Bartlett Publishers, USA.
Kruser, T.J. & Wheeler, D.L. 2010, "Mechanisms of resistance to HER family targeting antibodies", *Experimental cell research,* vol. 316, no. 7, pp. 1083-1100.
Kurtzman, A.L., Govindarajan, S., Vahle, K., Jones, J.T., Heinrichs, V. & Patten, P.A. 2001, "Advances in directed protein evolution by recursive genetic recombination: applications to therapeutic proteins", *Current opinion in biotechnology,* vol. 12, no. 4, pp. 361-370.
Laitinen, O.H., Airenne, K.J., Hytonen, V.P., Peltomaa, E., Mahonen, A.J., Wirth, T., Lind, M.M., Makela, K.A., Toivanen, P.I., Schenkwein, D., Heikura, T., Nordlund, H.R., Kulomaa, M.S. & Yla-Herttuala, S. 2005, "A multipurpose vector system for the screening of libraries in bacteria, insect and mammalian cells and expression in vivo", *Nucleic acids research,* vol. 33, no. 4, pp. e42.
Laitinen, O.H., Hytonen, V.P., Nordlund, H.R. & Kulomaa, M.S. 2006, "Genetically engineered avidins and streptavidins", *Cellular and molecular life sciences: CMLS,* vol. 63, no. 24, pp. 2992-3017.

Laitinen, O.H., Nordlund, H.R., Hytonen, V.P. & Kulomaa, M.S. 2007, "Brave new (strept)avidins in biotechnology", *Trends in biotechnology,* vol. 25, no. 6, pp. 269-277.

Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. 1993, "*PROCHECK*: a program to check the stereochemical quality of protein structures", *Journal of Applied Crystallography,* vol. 26, pp. 283-291.

Leader, B., Baca, Q.J. & Golan, D.E. 2008, "Protein therapeutics: a summary and pharmacological classification", *Nature reviews. Drug discovery,* vol. 7, no. 1, pp. 21-39.

Lehtonen, J.V., Still, D.J., Rantanen, V.V., Ekholm, J., Bjorklund, D., Iftikhar, Z., Huhtala, M., Repo, S., Jussila, A., Jaakkola, J., Pentikainen, O., Nyronen, T., Salminen, T., Gyllenberg, M. & Johnson, M.S. 2004, "BODIL: a molecular modeling environment for structure-function analysis and drug design", *Journal of computer-aided molecular design,* vol. 18, no. 6, pp. 401-419.

Leppiniemi, J., Maatta, J.A., Hammaren, H., Soikkeli, M., Laitaoja, M., Janis, J., Kulomaa, M.S. & Hytonen, V.P. 2011, "Bifunctional avidin with covalently modifiable ligand binding site", *PloS one,* vol. 6, no. 1, pp. e16576.

Lewis Phillips, G.D., Li, G., Dugger, D.L., Crocker, L.M., Parsons, K.L., Mai, E., Blattler, W.A., Lambert, J.M., Chari, R.V., Lutz, R.J., Wong, W.L., Jacobson, F.S., Koeppen, H., Schwall, R.H., Kenkare-Mitra, S.R., Spencer, S.D. & Sliwkowski, M.X. 2008, "Targeting HER2-positive breast cancer with trastuzumab-DM1, an antibody-cytotoxic drug conjugate", *Cancer research,* vol. 68, no. 22, pp. 9280-9290.

Livnah, O., Bayer, A., Wilchek, M. & Sussman, J.L. 1993a, "The structure of the complex between avidin and the dye, 2-(4'-hydroxyazobenzene) benzoic acid (HABA)", *FEBS,* vol. 328165, pp. 165-168.

Livnah, O., Bayer, E.A., Wilchek, M. & Sussman, J.L. 1993b, "Three-dimensional structures of avidin and the avidin-biotin complex", *Proceedings of the National Academy of Sciences of the United States of Americ,* vol. 90297, pp. 5076-5080.

Lonberg, N. 2008, "Fully human antibodies from transgenic mouse and phage display platforms", *Current opinion in immunology,* vol. 20, no. 4, pp. 450-459.

Looger, L.L., Dwyer, M.A., Smith, J.J. & Hellinga, H.W. 2003, "Computational design of receptor and sensor proteins with novel functions", *Nature,* vol. 423, no. 6936, pp. 185-190.

Mattheakis, L.C., Bhatt, R.R. & Dower, W.J. 1994, "An in vitro polysome display system for identifying ligands from very large peptide libraries", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 91, no. 19, pp. 9022-9026.

Metz, S., Haas, A.K., Daub, K., Croasdale, R., Stracke, J., Lau, W., Georges, G., Josel, H.P., Dziadek, S., Hopfner, K.P., Lammens, A., Scheuer, W., Hoffmann, E., Mundigl, O. & Brinkmann, U. 2011, "Bispecific digoxigenin-binding antibodies for targeted payload delivery", *Proceedings of the National Academy of Sciences of the United States of America,* doi 10.1073/pnas.1018565108.

Michaelson, J.S., Demarest, S.J., Miller, B., Amatucci, A., Snyder, W.B., Wu, X., Huang, F., Phan, S., Gao, S., Doern, A., Farrington, G.K., Lugovskoy, A., Joseph, I., Bailly, V., Wang, X., Garber, E., Browning, J. & Glaser, S.M. 2009, "Anti-tumor activity of stability-engineered IgG-like bispecific antibodies targeting TRAIL-R2 and LTbetaR", *mAbs,* vol. 1, no. 2, pp. 128-141.

Minshull, J. & Stemmer, W.P. 1999, "Protein evolution by molecular breeding", *Current opinion in chemical biology,* vol. 3, no. 3, pp. 284-90.

Morrison, S.L., Johnson, M.J., Herzenberg, L.A. & Oi, V.T. 1984, "Chimeric human antibody molecules: mouse antigen-binding domains with human constant region domains", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 81, no. 21, pp. 6851-6855.

Myhre, S., Henning, P., Friedman, M., Stahl, S., Lindholm, L. & Magnusson, M.K. 2009, "Re-targeted adenovirus vectors with dual specificity; binding specificities conferred by two different Affibody molecules in the fiber", *Gene therapy,* vol. 16, no. 2, pp. 252-261.

Nahta, R., Yu, D., Hung, M.C., Hortobagyi, G.N. & Esteva, F.J. 2006, "Mechanisms of disease: understanding resistance to HER2-targeted therapy in human breast cancer", *Nature clinical practice.Oncology,* vol. 3, no. 5, pp. 269-280.

Nardone, E., Rosano, C., Santambrogio, P., Curnis, F., Corti, A., Magni, F., Siccardi, A.G., Paganelli, G., Losso, R., Apreda, B., Bolognesi, M., Sidoli, A. & Arosio, P. 1998, "Biochemical characterization and crystal structure of a recombinant hen avidin and its acidic mutant expressed in *Escherichia coli*", *European journal of biochemistry,* vol. 256, no. 2, pp. 453-60.

Nelson, P.N., Reynolds, G.M., Waldron, E.E., Ward, E., Giannopoulos, K. & Murray, P.G. 2000, "Monoclonal antibodies", *Molecular pathology : MP,* vol. 53, no. 3, pp. 111-117.

Neuberger, M.S. & Milstein, C. 1995, "Somatic hypermutation", *Current opinion in immunology,* vol. 7, no. 2, pp. 248-254.

Nicaise, M., Valerio-Lepiniec, M., Minard, P. & Desmadril, M. 2004, "Affinity transfer by CDR grafting on a nonimmunoglobulin scaffold", *Protein science,* vol. 13, no. 7, pp. 1882-1891.

Nord, K., Gunneriusson, E., Ringdahl, J., Stahl, S., Uhlen, M. & Nygren, P.A. 1997, "Binding proteins selected from combinatorial libraries of an alpha-helical bacterial receptor domain", *Nature biotechnology,* vol. 15, no. 8, pp. 772-777.

Nordlund, H.R., Hytonen, V.P., Horha, J., Maatta, J.A., White, D.J., Halling, K., Porkka, E., Slotte, J.P., Laitinen, O.H. & Kulomaa, M.S. 2005, "Tetravalent single chain avidin: From subunits to protein domains via circularly permuted avidins", *The Biochemical journal,* vol. 392, pp.485-491.

Nordlund, H.R., Laitinen, O.H., Hytönen, V.P., Uotila, S.T., Porkka, E. & Kulomaa, M.S. 2004, "Construction of a dual chain pseudotetrameric chicken avidin by combining two circularly permuted avidins", *Journal of biological chemistry,* vol. 279, no. 35, pp. 36715-36719.

Nordlund, H.R., Laitinen, O.H., Uotila, S.T., Nyholm, T., Hytönen, V.P., Slotte, J.P. & Kulomaa, M.S. 2003, "Enhancing the thermal stability of avidin. Introduction of disulfide bridges between subunit interfaces", *Journal of biological chemistry,* vol. 278, no. 422421427, pp. 2479-283.

Nuttall, S.D., Krishnan, U.V., Hattarki, M., De Gori, R., Irving, R.A. & Hudson, P.J. 2001, "Isolation of the new antigen receptor from wobbegong sharks, and use as a scaffold for the display of protein loop libraries", *Molecular immunology,* vol. 38, no. 4, pp. 313-326.

Nygren, P.A. 2008, "Alternative binding proteins: affibody binding proteins developed from a small three-helix bundle scaffold", *The FEBS journal,* vol. 275, no. 11, pp. 2668-2676.

Nygren, P.A. & Skerra, A. 2004, "Binding proteins from alternative scaffolds", *Journal of immunological methods,* vol. 290, no. 1-2, pp. 3-28.

Orlova, A., Magnusson, M., Eriksson, T.L., Nilsson, M., Larsson, B., Hoiden-Guthenberg, I., Widstrom, C., Carlsson, J., Tolmachev, V., Stahl, S. & Nilsson, F.Y. 2006, "Tumor imaging using a picomolar affinity HER2 binding affibody molecule", *Cancer research,* vol. 66, no. 8, pp. 4339-4348.

Pazy, Y., Eisenberg-Domovich, Y., Laitinen, O.H., Kulomaa, M.S., Bayer, E.A., Wilchek, M. & Livnah, O. 2003, "Dimer-tetramer transition between solution and crystalline states of streptavidin and avidin mutants", *The Journal of bacteriology,* vol. 185, no. 14, pp. 4050-4056.

Pazy, Y., Kulik, T., Bayer, E.A., Wilchek, M. & Livnah, O. 2002, "Ligand exchange between proteins. Exchange of biotin and biotin derivatives between avidin and streptavidin", *Journal of biological chemistry,* vol. 277, no. 34, pp. 30892-30900.

Phan, G.Q., Yang, J.C., Sherry, R.M., Hwu, P., Topalian, S.L., Schwartzentruber, D.J., Restifo, N.P., Haworth, L.R., Seipp, C.A., Freezer, L.J., Morton, K.E., Mavroukakis, S.A., Duray, P.H., Steinberg, S.M., Allison, J.P., Davis, T.A. & Rosenberg, S.A. 2003, "Cancer regression and autoimmunity induced by cytotoxic T lymphocyte-associated antigen 4 blockade in patients with metastatic melanoma", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 100, no. 14, pp. 8372-8377.

Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kale, L. & Schulten, K. 2005, "Scalable molecular dynamics with NAMD", *Journal of computational chemistry,* vol. 26, no. 16, pp. 1781-1802.

Pugliese, L., Coda, A., Malcovati, M. & Bolognesi, M. 1993, "Three-dimensional structure of the tetragonal crystal form of egg-white avidin in its functional complex with biotin at 2.7 A resolution", *Journal of molecular biologyl,* vol. 231, no. 3, pp. 698-710.

Reichmann, D., Rahat, O., Cohen, M., Neuvirth, H. & Schreiber, G. 2007, "The molecular architecture of protein-protein binding sites", *Current opinion in structural biology,* vol. 17, no. 1, pp. 67-76.

Repo, S., Paldanius, T.A., Hytonen, V.P., Nyholm, T.K., Halling, K.K., Huuskonen, J., Pentikainen, O.T., Rissanen, K., Slotte, J.P., Airenne, T.T., Salminen, T.A., Kulomaa, M.S. & Johnson, M.S. 2006, "Binding Properties of HABA-Type Azo Derivatives to Avidin and Avidin-Related Protein 4", *Chemistry & biology,* vol. 13, no. 10, pp. 1029-1039.

Reynaud, C.A., Anquez, V., Grimal, H. & Weill, J.C. 1987, "A hyperconversion mechanism generates the chicken light chain preimmune repertoire", *Cell,* vol. 48, no. 3, pp. 379-388.

Riihimaki, T.A., Hiltunen, S., Rangl, M., Nordlund, H.R., Maatta, J.A., Ebner, A., Hinterdorfer, P., Kulomaa, M.S., Takkinen, K. & Hytonen, V.P. 2011, "Modification of the loops in the ligand-binding site turns avidin into a steroid-binding protein", *BMC biotechnology,* vol. 11, no. 1, pp. 64.

Rojas, N.R., Kamtekar, S., Simons, C.T., McLean, J.E., Vogel, K.M., Spiro, T.G., Farid, R.S. & Hecht, M.H. 1997, "De novo heme proteins from designed combinatorial libraries", *Protein science,* vol. 6, no. 12, pp. 2512-2524.

Ronnmark, J., Hansson, M., Nguyen, T., Uhlen, M., Robert, A., Stahl, S. & Nygren, P.A. 2002, "Construction and characterization of affibody-Fc chimeras produced in Escherichia coli", *Journal of immunological methods,* vol. 261, no. 1-2, pp. 199-211.

Rosano, C., Arosio, P. & Bolognesi, M. 1999, "The X-ray three-dimensional structure of avidin", *Biomolecular engineering,* vol. 16, no. 1-4, pp. 5-12.

Saven, J.G. 2011, "Computational protein design: engineering molecular diversity, nonnatural enzymes, nonbiological cofactor complexes, and membrane proteins", *Current opinion in chemical biology,* vol.15, no. 3, 452-457.

Schlehuber, S., Beste, G. & Skerra, A. 2000, "A novel type of receptor protein, based on the lipocalin scaffold, with specificity for digoxigenin", *Journal of molecular biology,* vol. 297, no. 5, pp. 1105-120.

Schlehuber, S. & Skerra, A. 2005, "Lipocalins in drug discovery: from natural ligand-binding proteins to "anticalins"", *Drug discovery today,* vol. 10, no. 1, pp. 23-33.

Schlehuber, S. & Skerra, A. 2002, "Tuning ligand affinity, specificity, and folding stability of an engineered lipocalin variant -- a so-called 'anticalin' -- using a molecular random approach", *Biophysical chemistry,* vol. 96, no. 2-3, pp. 213-228.

Schlehuber, S. & Skerra, A. 2001, "Duocalins: engineered ligand-binding proteins with dual specificity derived from the lipocalin fold", *Biological chemistry,* vol. 382, no. 9, pp. 1335-1342.

Schneider, S., Buchert, M., Georgiev, O., Catimel, B., Halford, M., Stacker, S.A., Baechi, T., Moelling, K. & Hovens, C.M. 1999, "Mutagenesis and selection of PDZ domains that bind new protein targets", *Nature biotechnology,* vol. 17, no. 2, pp. 170-175.

Schonfeld, D., Matschiner, G., Chatwell, L., Trentmann, S., Gille, H., Hulsmeyer, M., Brown, N., Kaye, P.M., Schlehuber, S., Hohlbaum, A.M. & Skerra, A. 2009, "An engineered lipocalin specific for CTLA-4 reveals a combining site with structural and conformational features similar to antibodies", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 106, no. 20, pp. 8198-8203.

Schroeder, H.W.,Jr & Cavacini, L. 2010, "Structure and function of immunoglobulins", *The Journal of allergy and clinical immunology,* vol. 125, no. 2 Suppl 2, pp. S41-52.

Segal, D.M., Weiner, G.J. & Weiner, L.M. 2001, "Introduction: bispecific antibodies", *Journal of immunological methods,* vol. 248, no. 1-2, pp. 1-6.

Shao, Z., Zhao, H., Giver, L. & Arnold, F.H. 1998, "Random-priming in vitro recombination: an effective tool for directed evolution", *Nucleic acids research,* vol. 26, no. 2, pp. 681-683.

Sharp, P.M., Cowe, E., Higgins, D.G., Shields, D.C., Wolfe, K.H. & Wright, F. 1988, "Codon usage patterns in *Escherichia coli, Bacillus subtilis, Saccharomyces cerevisiae, Schizosaccharomyces pombe, Drosophila melanogaster* and *Homo sapiens*; a review of the considerable within-species diversity", *Nucleic acids research,* vol. 16, no. 17, pp. 8207-8211.

Sidhu, S.S., Weiss, G.A. & Wells, J.A. 2000, "High copy display of large proteins on phage for functional selections", *Journal of molecular biology,* vol. 296, no. 2, pp. 487-495.

Skerra, A. 2008, "Alternative binding proteins: anticalins - harnessing the structural plasticity of the lipocalin ligand pocket to engineer novel binding activities", *The FEBS journal,* vol. 275, no. 11, pp. 2677-2683.

Skerra, A. 2007, "Alternative non-antibody scaffolds for molecular recognition", *Current opinion in biotechnology,* vol. 18, no. 4, pp. 295-304.

Skerra, A. 2003, "Imitating the humoral immune response", *Current opinion in chemical biology,* vol. 7, no. 6, pp. 683-693.
Skerra, A. 2001, "'Anticalins': a new class of engineered ligand-binding proteins with antibody-like properties", *Journal of biotechnology,* vol. 74, no. 4, pp. 257-275.
Smith, B.A. & Hecht, M.H. 2011, "Novel proteins: from fold to function", *Current opinion in chemical biology,* vol. 15, nro. 3, 421-426.
Smith, G.P. 1985, "Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface", *Science,* vol. 228, no. 4705, pp. 1315-1317.
Smith, J., Kontermann, R.E., Embleton, J. & Kumar, S. 2005, "Antibody phage display technologies with special reference to angiogenesis", *The FASEB journal : the Federation of American Societies for Experimental Biology,* vol. 19, no. 3, pp. 331-341.
Stemmer, W.P. 1994a, "DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 91, no. 22, pp. 10747-1051.
Stemmer, W.P. 1994b, "Rapid evolution of a protein in vitro by DNA shuffling", *Nature,* vol. 370, no. 6488, pp. 389-391.
Todorovska, A., Roovers, R.C., Dolezal, O., Kortt, A.A., Hoogenboom, H.R. & Hudson, P.J. 2001, "Design and application of diabodies, triabodies and tetrabodies for cancer targeting", *Journal of immunological methods,* vol. 248, no. 1-2, pp. 47-66.
Tonikian, R., Zhang, Y., Boone, C. & Sidhu, S.S. 2007, "Identifying specificity profiles for peptide recognition modules from phage-displayed peptide libraries", *Nature protocols,* vol. 2, no. 6, pp. 1368-1386.
Uchiyama, F., Tanaka, Y., Minari, Y. & Tokui, N. 2005, "Designing scaffolds of peptides for phage display libraries", *Journal of bioscience and bioengineering,* vol. 99, no. 5, pp. 448-456.
Underdown, B.J. & Schiff, J.M. 1986, "Immunoglobulin A: strategic defense initiative at the mucosal surface", *Annual review of immunology,* vol. 4, pp. 389-417.
Vanhercke, T., Ampe, C., Tirry, L. & Denolf, P. 2005, "Reducing mutational bias in random protein libraries", *Analytical biochemistry,* vol. 339, no. 1, pp. 9-14.
Villar, H.O. & Kauvar, L.M. 1994, "Amino acid preferences at protein binding sites", *FEBS letters,* vol. 349, no. 1, pp. 125-130.
Vincke, C., Loris, R., Saerens, D., Martinez-Rodriguez, S., Muyldermans, S. & Conrath, K. 2009, "General strategy to humanize a camelid single-domain antibody and identification of a universal humanized nanobody scaffold", *The Journal of biological chemistry,* vol. 284, no. 5, pp. 3273-3284.
Vriend, G. 1990, "WHAT IF: a molecular modeling and drug design program", *J Mol Graph,* vol. 8, no. 1, pp. 52-6, 29.
Walsh, G. 2005, "Biopharmaceuticals: recent approvals and likely directions", *Trends in biotechnology,* vol. 23, no. 11, pp. 553-558.
Weber, P.C., Cox, M.J., Salemme, F.R. & Ohlendorf, D.H. 1987, "Crystallographic data for Streptomyces avidinii streptavidin", *The Journal of biological chemistry,* vol. 262, no. 26, pp. 12728-12729.
Weng, S., Gu, K., Hammond, P.W., Lohse, P., Rise, C., Wagner, R.W., Wright, M.C. & Kuimelis, R.G. 2002, "Generating addressable protein microarrays with

PROfusion covalent mRNA-protein fusion technology", *Proteomics,* vol. 2, no. 1, pp. 48-57.

Wikman, M., Rowcliffe, E., Friedman, M., Henning, P., Lindholm, L., Olofsson, S. & Stahl, S. 2006, "Selection and characterization of an HIV-1 gp120-binding affibody ligand", *Biotechnology and applied biochemistry,* vol. 45, no. Pt 2, pp. 93-105.

Wikman, M., Steffen, A.C., Gunneriusson, E., Tolmachev, V., Adams, G.P., Carlsson, J. & Stahl, S. 2004, "Selection and characterization of HER2/neu-binding affibody ligands", *Protein engineering, design & selection : PEDS,* vol. 17, no. 5, pp. 455-462.

Wildling, L., Hinterdorfer, P., Kusche-Vihrog, K., Treffner, Y. & Oberleithner, H. 2009, "Aldosterone receptor sites on plasma membrane of human vascular endothelium detected by a mechanical nanosensor", *Pflugers Archiv : European journal of physiology,* vol. 458, no. 2, pp. 223-230.

Williams, A. & Baird, L.G. 2003, "DX-88 and HAE: a developmental perspective", *Transfusion and apheresis science : official journal of the World Apheresis Association : official journal of the European Society for Haemapheresis,* vol. 29, no. 3, pp. 255-258.

Winter, G., Griffiths, A.D., Hawkins, R.E. & Hoogenboom, H.R. 1994, "Making antibodies by phage display technology", *Annual review of immunology,* vol. 12, pp. 433-455.

Xu, J.L. & Davis, M.M. 2000, "Diversity in the CDR3 region of V(H) is sufficient for most antibody specificities", *Immunity,* vol. 13, no. 1, pp. 37-45.

Xu, L., Aha, P., Gu, K., Kuimelis, R.G., Kurz, M., Lam, T., Lim, A.C., Liu, H., Lohse, P.A., Sun, L., Weng, S., Wagner, R.W. & Lipovsek, D. 2002, "Directed evolution of high-affinity antibody mimics using mRNA display", *Chemistry & biology,* vol. 9, no. 8, pp. 933-942.

Zemlin, M., Klinger, M., Link, J., Zemlin, C., Bauer, K., Engler, J.A., Schroeder, H.W.,Jr & Kirkham, P.M. 2003, "Expressed murine and human CDR-H3 intervals of equal length exhibit distinct repertoires that differ in their amino acid composition and predicted range of structures", *Journal of molecular biology,* vol. 334, no. 4, pp. 733-749.

Zhao, H., Giver, L., Shao, Z., Affholter, J.A. & Arnold, F.H. 1998, "Molecular evolution by staggered extension process (StEP) in vitro recombination", *Nature biotechnology,* vol. 16, no. 3, pp. 258-261.

Chemistry & Biology 13, 1029–1039, October 2006 DOI 10.1016/j.chembiol.2006.08.006

Binding Properties of HABA-Type Azo Derivatives to Avidin and Avidin-Related Protein 4

Susanna Repo,[1] Tiina A. Paldanius,[2,4] Vesa P. Hytönen,[2,5] Thomas K.M. Nyholm,[1] Katrin K. Halling,[1] Juhani Huuskonen,[3] Olli T. Pentikäinen,[1,6] Kari Rissanen,[3] J. Peter Slotte,[1] Tomi T. Airenne,[1] Tiina A. Salminen,[1] Markku S. Kulomaa,[2,4] and Mark S. Johnson[1,*]

[1] Department of Biochemistry and Pharmacy
Åbo Akademi University
Tykistökatu 6
FI-20520 Turku
Finland

[2] NanoScience Center
Department of Biological and Environmental Science

[3] NanoScience Center
Department of Chemistry
P.O. Box 35
FI-40014 University of Jyväskylä
Finland

Summary

The chicken genome encodes several biotin-binding proteins, including avidin and avidin-related protein 4 (AVR4). In addition to *D*-biotin, avidin binds an azo dye compound, 4-hydroxyazobenzene-2-carboxylic acid (HABA), but the HABA-binding properties of AVR4 are not yet known. Differential scanning calorimetry, UV/visible spectroscopy, and molecular modeling were used to analyze the binding of 15 azo molecules to avidin and AVR4. Significant differences are seen in azo compound preferences for the two proteins, emphasizing the importance of the loop between strands β3 and β4 for azo ligand recognition; information on these loops is provided by the high-resolution (1.5 Å) X-ray structure for avidin reported here. These results may be valuable in designing improved tools for avidin-based life science and nanobiotechnology applications.

Introduction

Avidin, a basic glycoprotein from egg white, and streptavidin, the bacterial relative of avidin, bind a small vitamin molecule, *D*-biotin, with extremely high affinity (K_D = 10^{-13} to 10^{-15} M) [1, 2]. The nature of the specific, strong protein-ligand interaction has been extensively studied in order to identify the structural determinants of the high-affinity binding [3–5] and to engineer novel features in the development of more sophisticated tools for biotechnology applications [6–8]. Although the three-dimensional (3D) structures of both avidin and streptavidin in complex with *D*-biotin have been solved [9, 10], understanding the details of the biotin-binding process remains challenging [11].

Avidin belongs to a gene family that contains several avidin-related genes (AVRs) [12, 13] whose biological function is unknown. Previously, we produced AVRs in recombinant forms in both *E. coli* [14] and insect cells by utilizing a baculovirus expression system [15, 16]. AVRs possess fascinating properties: despite their lower biotin-binding affinity compared with avidin [15, 16], some AVRs have higher thermal stability than avidin [16, 17]. The 3D structures of AVR2 [17] and AVR4 [18] have been solved to high resolution.

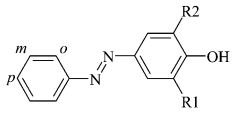
Avidin has a moderate level of affinity for a small molecule, an azo dye called HABA (4-hydroxyazobenzene-2-carboxylic acid, also called 2-(4′-hydroxybenzene) azobenzoic acid; see the structure in Table 1) [19]. Avidin-HABA binding is accompanied by a change in the spectral properties of the dye from yellow to red, a property used to measure the number of biotin-binding sites of avidin derivatives; the reaction is easily reversed by *D*-biotin. The spectroscopic features of HABA in avidin-based applications would obviate the need for radioactive labels; however, because the dissociation constant of the avidin-HABA interaction is only 6×10^{-6} M [19], HABA derivatives with higher affinity for avidin are needed for improved sensitivity. The 3D structure of the avidin-HABA complex has been solved: HABA binds to avidin as a hydrazone tautomer, planarity is thus lost, and an intramolecular hydrogen bond is formed [20]. Streptavidin also binds HABA as the hydrazone tautomer [21], although the affinity is lower (1×10^{-4} M) than that of avidin [2]. The 3D structure of the streptavidin-HABA complex has been solved [21].

Biotin binding to (strept)avidin leads to changes in the conformation [9, 10], stability [22], and rigidity [23] of the proteins. The most flexible part of avidin that interacts with biotin is a loop located between β strands 3 and 4 (L3,4 loop). In the X-ray structure of apo-avidin, the L3,4 loop is disordered [24]. The electron density maps of the avidin-biotin complex structure are clearly interpretable, and the L3,4 loop has a defined conformation [10]. Similarly for streptavidin, the corresponding loop is disordered in the absence of ligand [25]. Two states of the L3,4 loop, "open" and "closed," that are partially determined by the properties of the bound ligand have been observed in streptavidin [25]. In AVR4, the L3,4 loop adopts nearly identical conformations with and without bound biotin [18]. The L3,4 loop is disordered in the avidin-HABA structure [20], while the L3,4 loop adopts the "closed" conformation in structures of streptavidin complexed with HABA and HABA derivatives [21, 26].

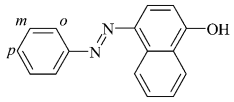
We have used various techniques to improve our understanding of the binding process of azo compounds to avidin and AVR4. We synthesized a set of 15 azo compounds (see Experimental Procedures and Supplemental Data [available with this article online]) and studied their interactions with the biotin-binding proteins by using differential scanning calorimetry

*Correspondence: johnson4@abo.fi
[4] Present address: Institute of Medical Technology, FI-33014 University of Tampere, Finland.
[5] Present address: Department of Materials, ETH Zurich, CH-8093 Zürich, Switzerland.
[6] Present address: Department of Biological and Environmental Science, P.O. Box 35, FI-40014 University of Jyväskylä, Finland.

Table 1. Functional Groups of the Synthesized Ligands



Compound	-COO$^-$ Position	R1	R2
HABA	*ortho*	H	H
1b	*meta*	H	H
1c	*para*	H	H
2b	*meta*	H	NO_2
2c	*para*	H	NO_2
4a	*ortho*	CH_3	CH_3
4b	*meta*	CH_3	CH_3
4c	*para*	CH_3	CH_3
5a	*ortho*	H	OH
5b	*meta*	H	OH
5c	*para*	H	OH
6b	*meta*	OH	CH_3
6c	*para*	OH	CH_3



Compound	-COO$^-$ Position
3a	*ortho*
3b	*meta*
3c	*para*

(DSC), UV/visible spectroscopy, and molecular modeling. Moreover, to our knowledge, we present the first high (1.48 Å)-resolution X-ray structure of chicken avidin.

Results

High-Resolution Crystal Structure of Avidin

To our knowledge, we report the first high-resolution X-ray structure of avidin at 1.48 Å resolution. The structure-determination statistics are in Table 2. The fold of the two monomers in the asymmetric unit and tetrameric assembly are nearly identical to the known avidin structures [20, 24, 27–31]: eight antiparallel β strands form a classic β barrel, with one open end serving as the biotin-binding site.

In contrast to other avidin structures, both the open and closed conformations of the L3,4 loop are observed in the electron density maps of this high-resolution structure (Figure 1A). In the closed conformation, the 12 residue-long L3,4 loop seals the ligand-binding pocket similarly to one of the avidin-biotin complexes (PDB code 1AVD [31]). In the 1.48 Å resolution structure, all residues within the L3,4 loop of the closed conformation could be traced in the electron density maps. The open conformation of the L3,4 loop diverges from the closed loop conformation at I34 and reunites with the closed loop conformation at K45. V37–S41 of the open loop conformation could not be traced through the electron density map. Thus, T35–A36, at the start of the L3,4 loop, and N42–K45, at the loop's end, could be built into the electron density of the open loop conformation (Figure 2A).

Before data collection, an avidin crystal was soaked with compound **3a**. Despite the color change observed after the soaking experiment, compound **3a** was not observed in the solved structure, but, instead, two glycerol molecules are in the ligand-binding pocket. The lack of detection of compound **3a** in the binding pocket may be explained by the low occupancy of the ligand in the binding sites. Thus, the ligand might be "invisible" in the crystal structure even though it might have bound to some protein molecules in the soaking experiment.

Table 2. Structure Determination Statistics for Avidin

Data Collection[a]	
Wavelength (Å)	0.804
Beamline	X13
Detector	CCD
Resolution (Å)	20–1.48 (1.58–1.48)
Unique observations	41,811 (7320)
I/σ	10.4 (2.9)
R factor[b] (%)	7.3 (47.0)
Completeness	99.3 (99.1)
Redundancy	4.0 (4.1)
Refinement	
Space group	$P2_12_12$
Unit cell	
a, b, c (Å)	72.9, 78.8, 43.0
α, β, γ (°)	90, 90, 90
Monomers per asymmetric unit	2
Resolution (Å)	20–1.48
R_{work} (%)	16.5
R_{free} (%)	19.0
Protein atoms	2,066
Heterogen atoms	36
Solvent atoms	149
Rmsd	
Bond lengths (Å)	0.013
Bond angles (°)	1.5
Ramachandran plot	
Residues in most favored regions	92.5%
Residues in additional allowed regions	7.5%

The PDB code for avidin is 1VYO.
[a] The numbers in parentheses refer to the highest-resolution bin.
[b] Observed R factor was taken from XDS [38].

DSC Analysis

The effect of the azo compounds on the stability of avidin and AVR4 was studied by using DSC (Table 3). Since increases in protein thermal stability upon ligand binding depends on the binding affinity, we have calculated the apparent binding constants at the temperature of protein unfolding, $K_b(T_m)$. Calculations are based on the temperature of protein unfolding in the absence and presence of ligand as well as on the change in enthalpy and heat capacity upon unfolding of the protein in the absence of ligand [32]. HABA, $K_D = 7 \times 10^{-6}$ M at pH 7 at room temperature [33], and 2,6-ANS, $K_D = 2 \times 10^{-4}$ M [34], which have moderate binding affinity for chicken avidin, were used as the control ligands. We also measured the stabilizing effect of the extreme affinity ligand, *D*-biotin, for avidin ($K_D \approx 10^{-15}$ M) and AVR4 ($K_D \approx 10^{-14}$ M) [1, 16].

Of the 18 ligand molecules that were analyzed, biotin produced the largest increase in avidin stability, $\Delta T_m = 34.5°C$ with $K_b(T_m)$ (M^{-1}) $= 3.1 \times 10^{11}$, and in AVR4 increased stability $\Delta T_m = 17.2°C$ with $K_b(T_m)$

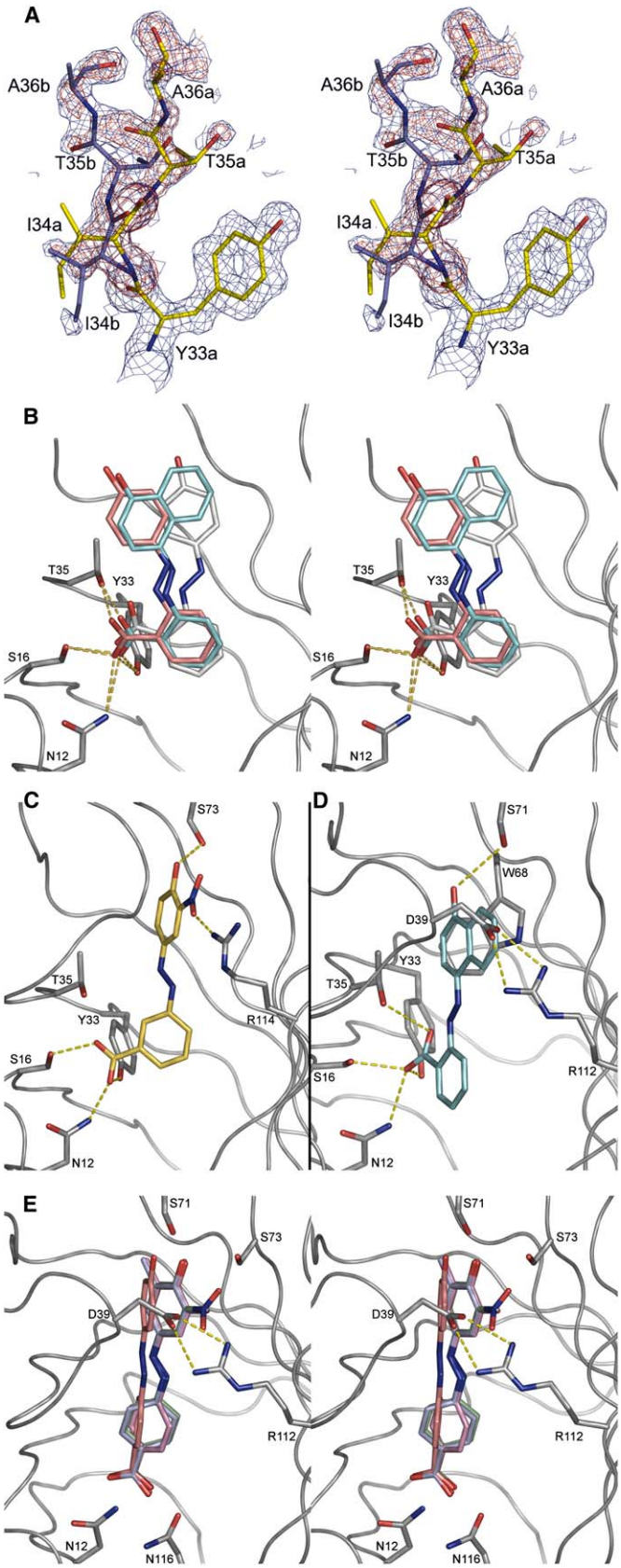


Figure 1. The L3,4 Loop and the Suggested Binding Mode of Some Azo Compounds

(A) The weighted $2F_o - F_c$ (blue) and $F_o - F_c$ (red) electron density maps of avidin showing the branch site for the two alternative L3,4 loops. The maps were calculated in the absence of the amino acids of the L3,4 loop (residues 34–45) and are shown here (a 2.2 Å radius around the atoms) along with the residues Y33, I34, T35, and A36 of the final structure of avidin (PDB code 1VYO). Contours are shown at 1.0σ and 3.0σ for the $2F_o - F_c$ and $F_o - F_c$ maps, respectively. Carbon atoms of residues from the "closed" (a) and "open" (b) conformation of the L3,4 loop are colored yellow and blue, respectively.

(B–E) The secondary structure of the protein in question is shown as a gray coil. Amino acids are represented as sticks; carbon atoms are shown in gray, oxygen atoms are shown in red, and nitrogen atoms are shown in blue. Putative hydrogen bonds are shown as dashed, yellow lines. (B) The suggested binding mode of HABA (orange) and **3a** (cyan) to avidin with the open L3,4 loop conformation. HABA (with light-gray carbon atoms) was also docked into the closed loop conformation of avidin. (C) The suggested binding mode of **2b** (yellow) to avidin. (D) The suggested binding mode of **3a** (cyan) to AVR4. (E) The suggested binding mode of **1c** (orange), **2c** (green), **4c** (purple), and **6c** (pink) to AVR4. For details on the suggested binding modes, see the main text.

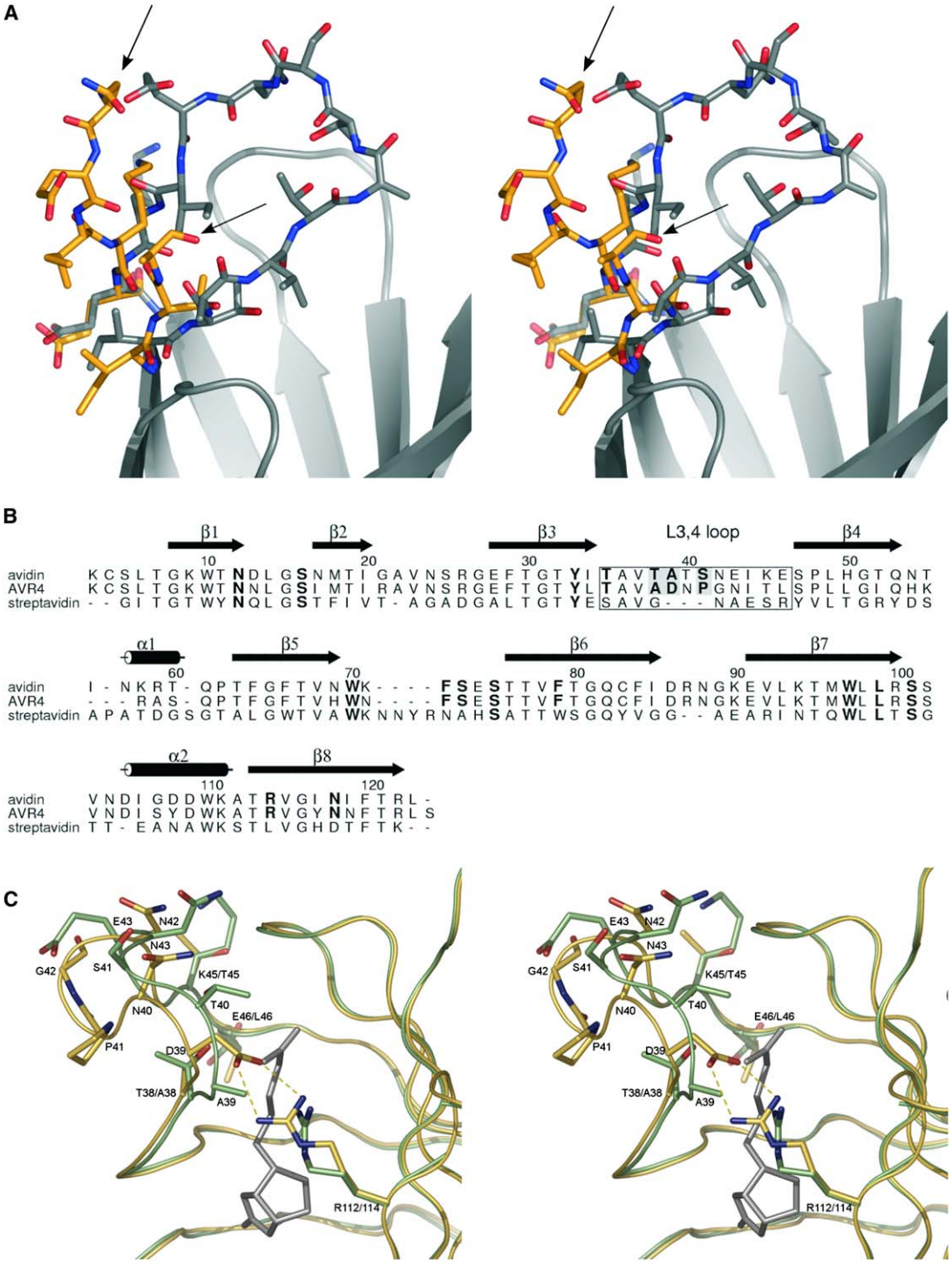


Figure 2. Two Conformations of the L3,4 Loop in Avidin, Structure-Based Sequence Alignment, and Comparison of the L3,4 Loop in Avidin and AVR4

(A) The high-resolution crystal structure of avidin (PDB code 1VYO). The secondary structure is shown in gray, and the amino acids in the L3,4 loop (I34–E46) are shown as sticks. The closed L3,4 loop conformation is in gray, and the open conformation is in orange. The arrows indicate the ends of the peptide chain in the open loop conformation.

(B) The structure-based sequence alignment of avidin (PDB code 1VYO), AVR4 (PDB code 1Y55), and streptavidin (PDB code 1SWE). The secondary structure elements of avidin are indicated, and the L3,4 loop of avidin, AVR4, and streptavidin is boxed. The key residues for azo compound binding in avidin and AVR4 are in bold (and when conserved, they are also in bold for streptavidin). The shaded residues indicate the most important differences (for details, see the text) between avidin and AVR4 sequences in the L3,4 loop. The alignment was performed with Vertaa implemented in Bodil [51], and the picture was produced with Alscript [54].

(C) The L3,4 loop of avidin (PDB code 1VYO, green) and of AVR4 (PDB code 1Y55, yellow). The structures were superimposed with Vertaa in Bodil [51]. Biotin, docked into the 1VYO structure with GOLD 2.2 [49], is shown in gray. Amino acids of the L3,4 loop, together with a conserved arginine residue from β8, are labeled and shown as sticks.

Table 3. The Transition Melting Temperature, T_m, of Avidin and AVR4 Determined in the Presence of Different Ligands

Protein:	Avidin			AVR4		
Ligand	T_m (°C)	ΔT_m (°C)	$K_b(T_m)$[a] (M^{-1})	T_m (°C)	ΔT_m (°C)	$K_b(T_m)$ (M^{-1})
—	82.5	—	ND[b]	109.9	—	ND[b]
HABA	92.8	10.3	2.9×10^5	110.9	1.0	3.2×10^3
1b	91.3	8.8	1.8×10^5	110.9	1.0	3.2×10^3
1c	86.3	3.8	2.0×10^4	114.9	5.0	4.2×10^4
2b	95.3	12.8	1.0×10^6	113.9	4.0	2.6×10^4
2c	87.1	4.6	2.6×10^4	115.6	5.7	5.8×10^4
3a	98.4	15.9	5.0×10^6	116.9	7.0	1.0×10^5
3b	91.5	9.0	1.8×10^5	111.1	1.2	4.0×10^3
3c	93.0	10.5	3.6×10^5	113.0	3.1	1.6×10^4
4a	95.6	13.1	1.2×10^6	113.6	3.7	2.2×10^4
4b	87.9	5.4	3.8×10^4	109.8	−0.1	ND[b]
4c	84.3	1.8	5.4×10^3	114.9	5.0	4.2×10^4
5a	94.5	12.0	7.0×10^5	112.2	2.3	9.8×10^3
5b	90.4	7.9	1.1×10^5	112.1	2.2	9.2×10^3
5c	86.5	4.0	2.0×10^4	114.5	4.6	3.5×10^4
6b	86.5	4.0	2.1×10^4	111.5	1.6	5.9×10^3
6c	84.1	1.6	4.9×10^3	115.2	5.3	4.8×10^4
2,6-ANS	84.6	2.1	6.8×10^3	114.8	4.9	4.0×10^4
D-biotin	117.0[c]	34.5	3.1×10^{11}	127.1	17.2	1.3×10^7

[a] The apparent binding constant (K_b) is calculated from data obtained at the T_m as described in Brandts et al. [32].
[b] Not determinable.
[c] From Hytönen et al. [16].

(M^{-1}) = 1.3×10^7. HABA significantly increased avidin stability (ΔT_m = 10.3°C, and $K_b[T_m]$ [M^{-1}] = 2.9×10^5), as did the five other azo compounds (ΔT_m > +10.0°C). Three molecules are *ortho*-derivatives of HABA (**3a**, **4a**, and **5a**), one is a *meta*-derivative (**2b**), and one is a *para*-derivative (**3c**). Overall, the stabilizing effect of the *ortho*-derivatives of HABA was largest: on average, ΔT_m = 12.8°C and $K_b(T_m)$ (M^{-1}) = 1.8×10^6. The *meta*-derivatives showed somewhat poorer stabilization of avidin (average ΔT_m = 8.0°C; $K_b[T_m]$ [M^{-1}] = 2.5×10^5), while the *para*-derivatives were the least effective in terms of stabilizing the avidin complex (average ΔT_m = 4.4°C; $K_b[T_m]$ [M^{-1}] = 7.3×10^4). The control ligand, 2,6-ANS, stabilized avidin rather poorly, with ΔT_m = 2.1°C and $K_b(T_m)$ (M^{-1}) = 6.8×10^3.

In comparison to avidin, AVR4 showed somewhat different behavior in the presence of the azo compounds. HABA was relatively inefficient in stabilizing AVR4 (ΔT_m = 1.0°C and $K_b[T_m]$ [M^{-1}] = 3.2×10^3), as were the other *ortho*-derivatives (average ΔT_m = 3.5°C; $K_b[T_m]$ [M^{-1}] = 3.4×10^4), in contrast to the *para*-derivatives with, on average, a ΔT_m of 4.8°C and $K_b(T_m)$ (M^{-1}) = 4.0×10^4. The *meta*-derivatives were the least effective in stabilizing AVR4: with, on average, a ΔT_m = 1.7°C and $K_b(T_m)$ (M^{-1}) = 9.7×10^3. Interestingly, 2,6-ANS stabilized AVR4 with a ΔT_m = 4.9°C and $K_b(T_m)$ (M^{-1}) = 4.0×10^4, similar to the average values for the *para*-derivatives.

Altogether, the DSC analyses suggest that avidin and AVR4 do not have similar binding affinities for the azo compounds. Avidin exhibits stronger binding for the *ortho*-forms, whereas AVR4 binds the *para*-forms more strongly. Furthermore, AVR4-2,6-ANS binding appears to be stronger than that of avidin.

UV/Visible Spectroscopy

The spectral characteristics of the azo compounds in aqueous solution were characterized by UV/visible spectroscopy. Between 300 and 650 nm, two peaks were found at 350.7 nm and 457.9 nm (average values) in the absorption spectrum for most compounds. Avidin is known to affect the absorption spectrum of HABA dramatically [33]. Consequently, we examined the spectral properties of HABA derivatives in the presence of avidin and AVR4. In general, changes in the absorbance spectrum reflect alterations in the physicochemical environment surrounding the ligand in the protein-bound form as well as any changes in double bond conjugation that result from binding. We mainly observed bathochromic shifts of the absorption maximum; however, in some instances, hypsochromic shifts of the absorption maximum were also observed. In the case of HABA, there are two peaks, at 346 nm and 439 nm in the absorbance spectrum. These peaks are due to the two tautomeric conformations of HABA; upon binding avidin, the conformation absorbing near 500 nm is favored. It has been shown that HABA bound to avidin forms the hydrazone tautomer [20]. Thus, there are two simultaneous changes that result from interactions with azo ligands: (1) the accumulation of a specific ligand conformation, which leads to absorption intensity differences between peaks; and (2) spectral shifts that are the result of inter- and intramolecular hydrogen bond formation and rearrangement of double bonds. Here, the spectral shifts are most informative, since they provide information on the relative strength of the avidin and AVR4 complexes with the azo compounds. The measured absorbance peak maxima (A_{max}) are listed in Table 4.

Similarly to HABA, the presence of avidin caused the longer-wavelength peak of the *ortho*-derivatives to move toward longer wavelengths (the average ΔA_{max} = 34 nm). This increase in the maximum absorption wavelength probably results from the formation of an intramolecular hydrogen bond in the *ortho*-derivatives of the azo compounds [33]. Such radical changes were not seen in the absorbance spectra of the *meta*- and

Table 4. Spectral Analysis of the Azo Molecules in the Absence and Presence of Avidin and AVR4

			+Avidin		+AVR4	
Ligand	$A_{max}(1)$[a]	$A_{max}(2)$	$A_{max}(1)$	$A_{max}(2)$	$A_{max}(1)$	$A_{max}(2)$
HABA	346	439	343	496	339	494
1b	348	431	343	432	332	413
1c	354	438	357	441	393	ND[b]
2b	371	419	364	431	351	422
2c	378	430	389	456	382	452
3a	325	493	329	522	335	518
3b	324	482	331[c]	499	330[d]	488
3c	327	481	340	511	357	519
4a	350	477	323	500	325	497
4b	354	444	349	444	343	447
4c	360	456	360	457	358	469
5a	360	474	ND[e]	501	ND[e]	499
5b	359	454	375	442	363	453
5c	367	461	362	470	ND[b]	479
6b	343	472	340	477	318	446
6c	347	474	347	474	343	470

[a] The measured spectra of the azo molecules was analyzed by fitting two Gaussian peaks to the data, and the absorption maxima (in nm) were determined from the fitted curves.
[b] Only one peak was observed in the spectrum.
[c] A new peak with A_{max} at 405 nm also appeared.
[d] A new peak with A_{max} at 411 nm also appeared.
[e] The lower-wavelength peak was negligible.

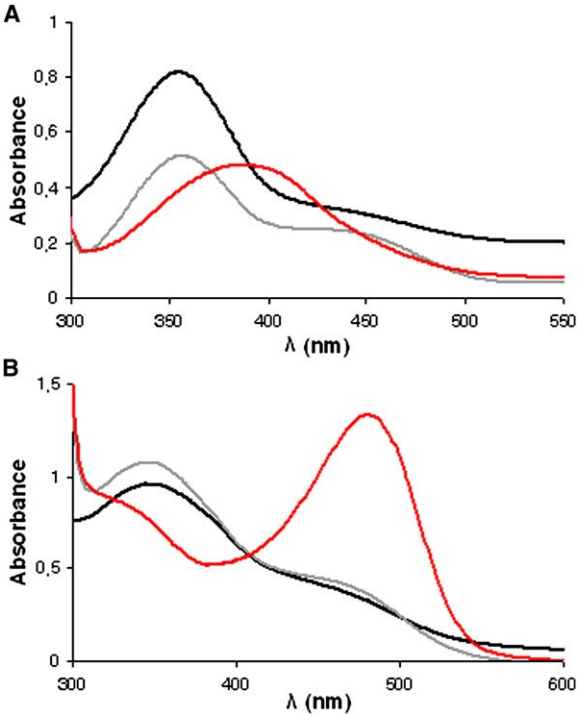


Figure 3. UV/Visible Spectra of **1c** and **5c**

(A) The UV/visible spectra of 4-hydroxyazobenzene-4-carboxylic acid (**1c**; black) with avidin (gray) and AVR4 (containing the mutation C122S; red).

(B) The UV/visible spectra of 3′,4-dihydroxyazobenzene-4-carboxylic acid (**5c**; black) in the presence of avidin (gray) and AVR4 (C122S; red).

the *para*-forms (average ΔA_{max} = 3.8 nm and 11.5 nm, respectively). Compounds **3a**, **3c**, **4a**, and **5a** with avidin result in the largest spectral changes, suggesting that the compounds have high affinity for avidin, while compounds **1c**, **3b**, and **5c** led to moderate spectral changes. Compounds **1b**, **4b**, **4c**, and **5b** with avidin produced only negligible changes in the spectra. The spectroscopic data for compounds **2b** and **2c** is unclear, and the affinity of these compounds for avidin cannot be inferred from the data.

Like avidin, the presence of AVR4 induced similar spectral changes in the maximum peak found at longer wavelengths for the *ortho*-compounds (average ΔA_{max} = 31.3 nm). For the *meta*-forms, AVR4 induced a slight decrease in A_{max} at longer wavelengths (average ΔA_{max} = −5.5 nm), while the A_{max} of the *para*-forms increased by 14.5 nm in the presence of AVR4. According to the spectroscopic studies, compounds **1c**, **3a**, **4a**, **5a**, and **5c** seem to produce the most significant spectral changes, which is indicative of high affinity for AVR4, while HABA and compound **3b** lead to moderate spectral changes; compounds **1b**, **2c**, **3c**, **4b**, **4c**, **5b**, **6b**, and **6c** produced only negligible changes to spectra in the presence of AVR4, suggesting that AVR4 has a low binding affinity for each compound.

In this analysis, the most interesting difference between avidin and AVR4 was found with compounds **1c** and **5c** (Figure 3). In the presence of avidin, the spectral changes of these compounds were negligible. The presence of AVR4, however, introduced a single peak centered at 393 nm in the case of **1c** (Figure 3A) and at 479 nm in the case of **5c** (Figure 3B). The induction of a single peak upon protein-ligand interaction suggests that a single tautomer of both **1c** and **5c** binds to AVR4 with higher affinity than for the corresponding avidin complexes. Based on the available data, however, we are not able to determine which tautomer of **1c** and **5c** is in question. Heavy computational calculations or high-resolution crystal structures would be needed to answer that question. With **3b**, a unique peak with an A_{max} of ~410 nm appeared with both avidin and AVR4. This peak emerges at a wavelength between the shorter (~300 nm)- and the longer (~500 nm)-wavelength peaks seen for **3b** and the other azo compounds. This is an interesting observation requiring further study. Overall, compared with avidin, AVR4 induced larger changes in the spectra of the *para*-derivatives of the azo compounds (average ΔA_{max} = 11.5 nm for avidin and 14.5 nm for AVR4).

Ligand Docking Studies

All 15 azo compounds that were synthesized (Table 1), as well as HABA and 2,6-ANS, were docked into 3 different structures: the avidin structure (PDB code 1VYO) in both the closed and open conformations and the AVR4 structure (PDB code 1Y55). The sequences of avidin and AVR4 are 77% identical [18], and the major differences that could affect ligand binding are located along the L3,4 loop (residues 35–46). The sequence alignment between avidin and AVR4 is shown in Figure 2B. The L3,4 loop in avidin and AVR4 is 12 amino acids long, but there are several large differences in the properties of residues forming the loop in each protein: a salt bridge is formed in AVR4 between D39 (A39 in avidin)

and R112 (R114 in avidin, see Figure 2C), and a proline residue at position 41 is S41 in avidin.

Altogether, the experimental studies identified five azo compounds with affinities for avidin that are similar to (**3c**) or larger (**2b**, **3a**, **4a**, and **5a**) than the affinity of HABA for avidin. According to the docking studies, the benzoate group of the three best *ortho*-compounds, **3a**, **4a**, and **5a**, would form interactions identical to those observed in the avidin-HABA complex structure [20]. An oxygen atom in the benzoate group of **3a**, **4a**, and **5a** is positioned so that it could hydrogen bond with S16 and T35 (unless specified otherwise, interactions are with the side chain of the indicated residue), while a second oxygen atom would form up to three hydrogen bonds with N12, S16, and Y33 (Figure 1B). In addition, the benzoate ring would interact with F79, W97, and W110 from the adjoining monomer. The docking studies indicate that the additional phenyl ring of **3a** fully occupies a pocket formed by W70, F72, S73, S75, and L99. Similarly, one of the methyl groups found in **4a** and the additional hydroxyl group of **5a** are also docked into this pocket (the hydroxyl group hydrogen bonding with S73), but the methyl and the hydroxyl groups on their own are not bulky enough to fill the pocket.

Compound **2b** is a *meta*-derivative of HABA and has an additional NO_2 group. As docked, the *meta*-carboxylate group would be buried deeper within the binding pocket of avidin compared to the *ortho*-derivative. According to the docking results, it appears likely that there is a hydrogen bond between one of the carboxylate oxygens of the ligand and S16, as well as two hydrogen bonds between the other carboxylate oxygen and N12 and Y33 (Figure 1C). Upon avidin binding, the NO_2 group of **2b** is most likely accommodated at a solvent-accessible site next to S75, S101, and R114, where the NO_2 group would be positioned to form hydrogen bonds with R114.

The *para*-compound **3c** has a similar affinity to avidin as HABA. The interactions of **3c**, at the bottom of the ligand-binding pocket, include possible hydrogen bonds with N12 and N118 and hydrophobic interactions with F79, W97, and W110 from the adjoining monomer. **3c** is longer than the *ortho*-compound **3a**; as a consequence, **3c** would extend further out of the binding pocket.

Experimental results show that 2,6-ANS binds to AVR4 with a significantly higher affinity than to avidin, in good agreement with our docking studies that showed that 2,6-ANS fits nicely into the binding pocket of AVR4 and that the phenyl ring of 2,6-ANS would stack with W68 and F70. In addition, the interaction of the SO_3 group of 2,6-ANS with AVR4 appears more favorable than with avidin: the SO_3 group could hydrogen bond to N12, S16, Y33, and N116 of AVR4. When docked to avidin, the phenyl ring of 2,6-ANS would clash with residues of the L3,4 loop, especially with T40.

Interestingly, the experimental studies indicate that all but two of the HABA derivatives (**1b** and **4b**) have affinities for AVR4 that are higher than the affinity of HABA for AVR4. The docking studies indicate that regardless of the closed conformation of the L3,4 loop of AVR4, the azo compound with the highest affinity for AVR4 (and for avidin), **3a**, can fit into the binding pocket of AVR4. The naphthalene group of **3a** is capable of forming π-π stacking interactions with W68. In addition, a hydrogen bond could form between S71 (Figure 1D) and the hydroxyl group of **3a**. Interactions with the benzoate group of **3a** in the docked complexes are identical to those observed for the avidin-HABA complex [20].

The *para*-derivatives of HABA, **1c**, **2c**, **4c**, **5c**, and **6c**, have very similar affinity for AVR4, and according to the docking studies, the binding mode of these compounds to AVR4 is also similar to each other. The benzoate group of the *para*-compounds is positioned where it could form hydrogen bonds with N12 and N116, at the bottom of the ligand-binding pocket of AVR4 (Figure 1E), as well as hydrophobic interactions with Y33, F77, W95, and W108 from the adjacent monomer. In addition, the phenolic part of the compounds is nicely docked between the hydrophobic side chains of A38, W68, F70, and L97, and a hydrogen bond can be positioned to link the phenolic hydroxyl group and the hydroxyl group of S71. The *para*-compound with the highest affinity for AVR4, **2c**, has an NO_2 group attached to the phenol ring. The docking studies show that this polar group could hydrogen bond with S73 and R112 (Figure 1E). The additional hydroxyl group of **6c** seems capable of hydrogen bonding with D39 and S73, while the additional methyl group is docked between the side chains of W68 and F70. Compounds **1c** and **4c** show equal affinity for AVR4. **1c** more closely resembles HABA in that no additional groups are attached to the phenol ring, whereas **4c** has two additional methyl groups on the ring. One of the methyl groups is docked between the side chains of W68 and F70, and the other methyl group is accommodated within a region formed by D39, S73, L97, and R112.

Discussion

In egg white, biotin binding is virtually irreversible and is needed to protect the growing embryo from infection by microbes. For applications in biotechnology, however, such irreversible binding is not always useful. While there is no need, and it is probably not even possible, to design any ligands with affinity better than that of *D*-biotin, research has focused on finding compounds with high, but reversible, affinity for avidin. The organic dye compound HABA (4-hydroxyazobenzene-2-carboxylic acid), with moderate affinity for avidin, is a good starting point for the design of novel avidin-binding compounds. Due to the spectroscopic features of HABA and its derivatives, their use in avidin-based applications would be safer than assays employing radioactive labels.

The formation of the complex between HABA and avidin is accompanied with a change in the spectral properties of HABA [33]. In the original paper, Green also reported that a *meta*-derivative of HABA bound to the avidin biotin-binding site, although the spectral changes were less dramatic. Similar spectral changes were observed for the complex of HABA and streptavidin, but the affinity of HABA for streptavidin (K_D = 100×10^{-6} M) is less than for avidin ($K_D = 6 \times 10^{-6}$ M) [19, 33]. Green has also reported the specific binding of other similar dye compounds to the biotin-binding site of avidin [33]. Much later, Weber et al. [26] successfully developed derivatives of HABA that show higher affinity toward streptavidin.

Avidin has been reported to have pseudocatalytic activity, meaning that avidin is capable of enhancing the hydrolysis of biotinyl ester derivatives, whereas streptavidin efficiently protects the same derivatives from hydrolysis [27]. This and other studies on the pseudocatalytic activity of avidin and streptavidin demonstrate the importance of the L3,4 loop as a molecular regulator in constraining the binding of biotin and biotin derivatives and in converting the protein into a pseudoenzyme [18, 35]. Comparisons of the crystal structures of streptavidin have shown that the L3,4 loop can adopt two structural states depending on the bound ligand [25]. In the case of avidin, the corresponding loop is disordered in the apo form [24], but biotin binding stabilizes the conformation of the loop and "locks" it into the closed state [10].

According to the experimental studies presented here, the *ortho*-derivatives of HABA bind better to avidin, whereas the *para*-derivatives bind better to AVR4. The avidin and AVR4 sequences are very similar throughout the secondary structure elements, except for the L3,4 loop (Figure 2B). Based on the docking studies, we suggest that the *ortho*-derivatives of HABA bind avidin with a conformation of the L3,4 loop that is more open. The results previously reported by Ellison et al. [36] also support this hypothesis. They have shown that HABA increases the affinity of proteinase K for avidin, and that, as a result, the L3,4 loop of avidin is cleaved. Furthermore, in the crystal structure of the avidin-HABA complex the L3,4 loop is not visible [20], suggesting that it is mobile. According to the docking results, if the benzoate group of HABA is docked into the closed loop conformation of avidin, it is not able to maintain hydrogen-bonding interactions with residues (N12, S16, T35, and Y33) along the "bottom" of the binding pocket (Figure 1B). Taken together, the current experimental data and docking studies support a model for HABA binding to avidin in which HABA increases the mobility of the L3,4 loop.

In the AVR4 structure, the conformation and flexibility of the L3,4 loop is constrained by the salt bridge formed between the side chains of D39 and R112 [18]. In contrast, the L3,4 loop of avidin is not constrained by a salt bridge, resulting in more conformational flexibility. Consequently, the *ortho*-compounds cannot be accommodated within the AVR4-binding pocket as well as in avidin. Due to the proline residue (P41) in the L3,4 loop of AVR4, the loop also becomes somewhat "shorter" than the L3,4 loop in avidin. As a result, the mouth of the binding pocket of AVR4 is wider than in the avidin structure (Figure 2C). Since the *para*-derivatives of HABA are more elongated than the *ortho*- and *meta*-derivatives, they extend further out from the binding pocket. In avidin, the *para*-derivatives collide with T40, but, in AVR4, there is sufficient space to accommodate the longer *para*-derivatives (Figure 2C). These docking studies are in good agreement with our DSC studies, which show that AVR4 binds the the *para*-compounds more strongly, while avidin has stronger binding affinity for the *ortho*-compounds.

The exception from the previous pattern is naphthyl-HABA, compound **3a**, which, according to the DSC studies, has the highest affinity for both avidin and AVR4. According to the docking studies, avidin would bind this *ortho*-compound with the open conformation of the L3,4 loop, in contrast to AVR4 with the added salt bridge that suggests binding might involve the closed loop conformation. The crystal structure of the streptavidin complex and compound **3a** shows that **3a** is bound to a closed loop conformation [26]. It seems that the shorter length of the L3,4 loop of streptavidin (Figure 2B) is the reason it adopts the closed loop conformation. Although the L3,4 loop of AVR4 also appears to be somewhat shorter than that of avidin (due to a kink in the loop introduced by P41), there are other features of the loop contributing to the mode of ligand binding. First of all, the salt bridge between the L3,4 loop and the β8 strand holds the loop tightly in its position. Second, in the L3,4 loop of AVR4, the methyl group of A38 would form a hydrophobic interaction with the two-ring system of compound **3a**. The corresponding residue in avidin is threonine (T38), which, due to its larger volume and polar hydroxyl group, hinders the binding of compound **3a** to the closed L3,4 loop conformation of avidin.

Significance

The avidin family of biotin-binding proteins is widely exploited in biotechnological applications and, more recently, in nanobiotechnology. There are obvious advantages to using a high-affinity ligand that alters its spectral characteristics when bound to an avidin protein, allowing one to measure spectrophotometrically the degree of successful complexation that has occurred. The organic azo dye, HABA (4-hydroxyazobenzene-2-carboxylic acid), does undergo such spectral changes when the avidin-HABA complex is formed, but the binding affinity of avidin for HABA is much less than the extremely tight binding that occurs with biotin. Thus, in order for a HABA-type molecule to be of use, the affinity must be substantially improved or another avidin-like molecule with tighter binding for a HABA compound should be identified. Here, we synthesized 15 derivatives of HABA, and we analyzed their binding characteristics with avidin and with a related protein, AVR4, by using DSC and UV/visible spectroscopy. We show that a number of these derivatives do have higher affinity for avidin, but that some compounds prefer to bind to AVR4. To our knowledge, we have solved the first high-resolution 1.5 Å structure of avidin that reveals novel structural details about the L3,4 loop, a loop that appears to be a major determinant defining the binding preferences of avidin and AVR4 for the HABA compounds. Molecular modeling was used to help position the compounds within the ligand-binding sites in the 3D structures of avidin and AVR4. These studies reveal the molecular basis for differences in ligand binding and provide firm details that can be used to determine how to proceed with the protein structure-based approach used to modify azo compounds to achieve higher affinity and selectivity for avidin or AVR4.

Experimental Procedures

Chemicals

2-aminobenzoic acid, 2-nitrophenol, and 1,2-dihydroxybenzene were purchased from Fluka (Buchs, Switzerland), and 4-aminobenzoic acid, 2,6-dimethylphenol, and 1-napthol were purchased from

Merck (Darmstadt, Germany). 1,2-dihydroxy-3-methylbenzene was purchased from Aldrich Chemical Company. 2-anilinonaphthalene-6-sulfonic acid (2,6-ANS) was from Invitrogen. *D*-biotin was purchased from Sigma, and chicken avidin was a generous gift from Belovo S.A., Bastogne, Belgium.

Synthesis and Characterization of Azo Molecules

HABA derivatives (for a table of the molecular structures, see Table 1) were synthesized by coupling diazotized aminobenzoic acid to the phenol and naphthol derivatives [26]. Ice-cooled aqueous $NaNO_2$ (7.3 mmol) was slowly poured into a stirred solution of *ortho-/meta-/para*-aminobenzoate (7.3 mmol) in hydrochloric acid (18%) while the temperature was kept below 5°C. The mixture containing the diazonium salt was cautiously poured into an ice-cooled aqueous solution of the phenol/naphthol derivatives (14.0 mmol) and NaOH (16.8 mmol). The reaction mixture was stirred at 5°C for 1 hr and neutralized with hydrochloric acid/sodium acetate. The resulting precipitates were collected, washed with water, and dried in vacuo. The product was recrystallized from the ethanol-water mixture and dried.

In the synthesis of 1,2-dihydroxyazobenzene carboxylic acid, 1,2,3-dihydroxyazobenzene carboxylic acid, and 1,2-dihydroxy-3-methylazobenzene carboxylic acid, the reactions were slightly different because the hydroxyl groups were protected by aluminium sulfate. Ice-cooled aqueous $NaNO_2$ (25 mmol) was poured into a stirred solution of *ortho-/meta-/para*-aminobenzoate (25 mmol) in hydrochloric acid (4.5%) while the temperature was kept below 5°C. 1,2-dihydroxybenzene/1,2,3-dihydroxybenzene/1,2-dihydroxy-3-methylbenzene (25 mmol) was diluted with aqueous $Al_2(SO_4)_3$ (14 mmol), and the diazonium salt solution was poured cautiously into this ice-cooled solution. The reaction mixture was stirred at 5°C for 40 min, and 25 ml CH_3COONa (20%) was added to the reaction mixture. The solution was acidified by adding 10 ml concentrated HCl, and the resulting reddish precipitates were collected, washed with water, and dried in vacuo. The products were recrystallized from an ethanol-water mixture and dried in vacuo.

The final products were confirmed by ^{1}H-NMR (Bruker Avance DPX 250 FT), ^{13}C-NMR (Bruker Avance DPX 500), and by MS. The ^{1}H-NMR spectra were recorded at 250 MHz, and ^{13}C spectra were recorded at 500 MHz in d-DMSO at 30°C. The detailed list of synthesized compounds with confirmed compound characteristics is in Supplemental Data.

Differential Scanning Calorimetry Analysis

The thermodynamics of the denaturation process of avidin isolated from chicken (Belovo S.A., Bastogne, Belgium) and the AVR4 protein produced in *E. coli* [14, 37], in the presence of different ligands (3:1 molar ratio), were studied by using a Nano II differential scanning calorimeter (Calorimetric Science Corporation, Provo, UT) as previously described [7]. AVR4 carried a mutation, C122S, that prevents intermolecular disulfide bridge formation without affecting either ligand binding or thermostability [16]. Sodium phosphate buffer (50 mM, pH 7.0) containing 100 mM NaCl was used in the measurements. Thermograms were analyzed with Origin 6.0 software. The apparent binding constant at the temperature of protein unfolding $K_b(T_m)$ was calculated as described in Brandts et al. [32] by using the following equation:

$$K_b(T_m) = [\exp[-\Delta H(T_0)/R(1/T_m - 1/T_0) + \Delta C_p/R(\ln T_m/T_0 + T_0/T_m - 1)] - 1]/[L]T_m, \quad (1)$$

where $\Delta H(T_0)$ is the change in enthalpy upon unfolding in the absence of ligand, R is the gas constant, T_m is the unfolding temperature in the presence of ligand, T_0 is the unfolding temperature in the absence of ligand, ΔC_p is the change in heat capacity upon unfolding, and $[L]T_m$ is the free ligand concentration at T_m. It is assumed that the unfolding process is a two-state transition and that the ligands bind only to the low-temperature, native conformation. In the calculation of the binding constants, we used the respective values for avidin and AVR4: $\Delta H(T_0)$, 329 kJ/mol and 460 kJ/mol; ΔC_p, 15.3 kJ/(K mol) and 9.1 kJ/(K mol) [16]. The value of ΔC_p used for AVR4 originates from measurements performed on AVR4 protein produced in insect cells and not in bacteria. Even a significant deviation in the value of ΔC_p would, however, only have a very small impact on the calculated $K_b(T_m)$s. Such a deviation would not affect the relative binding affinities for one protein and several different ligands.

UV/Visible Spectroscopy

The optical properties of the synthesized azo compounds were characterized with a PerkinElmer UV/visible spectrometer. Spectra were obtained for samples in 50 mM Na-phosphate buffer (pH 7.0) containing 100 mM NaCl. The azo compounds were diluted to a final concentration of 20 mM, and the spectrum for each compound was measured at wavelengths between 300 and 650 nm. Furthermore, the spectrum of each of the azo compounds (20 mM) in the presence of 1 mg/ml chicken avidin (Belovo S.A., Bastogne, Belgium) or AVR4 (produced in *E. coli* [14]) was measured. The spectrum for each compound was analyzed by using multipeak analysis (Microcal Origin 7.0), fitting two Gaussian peaks to the data.

Crystal Structure Determination

The recombinant chicken avidin protein used for structure determination with X-ray crystallography was produced in *E. coli*, purified, and crystallized as previously described [14]. Briefly, bar-like crystals were obtained at +22°C by using the hanging drop vapor diffusion method. Equal volumes (1 μl) of protein (0.5 mg/ml) in 50 mM Na acetate (pH 4) + 20 mM NaCl and well solution of 0.1 M MES (pH 6.6) + 24% PEG 8000 + 0.2 M Mg acetate were employed. Before data collection, a crystal was soaked with compound **3a** (1 mM in drop) for several hours at +22°C. Glycerol (23% v/v) was added to the crystallization drop to serve as a cryoprotectant just prior to flash freezing in a 100 K nitrogen stream.

Diffraction data were collected at 100 K at the synchrotron beamline X13 (EMBL-Hamburg) from a single crystal. The data were indexed, integrated, and scaled with the program package XDS [38]. The structure was solved with the molecular replacement technique and by applying programs from the CCP4i Suite [39]. An existing 2.7 Å avidin structure (PDB code 1AVD [31]) was used as a trial model in molecular replacement, which was carried out with the program AMoRe [40]. The X-ray structure was refined with Refmac5 [41] by using TLS [42] and was modified and rebuilt with the program O [43]. Solvent atoms were added to the model with the automatic procedure of ARP/wARP [44], whereas sulfate ions and glycerol molecules were added manually in O. The final model was analyzed with the programs PROCHECK [45] and WHATIF [46].

Docking Studies

The 3D structures of 2,6-ANS and the azo compounds (Table 1), including two different tautomers for each azo compound, were built and energy minimized in Sybyl (Tripos, St. Louis, MO). The Conjugate Gradient minimization method was used along with the MMFF94s force field and MMFF94 charges [47, 48]. The termination gradient was set to 0.05 kcal/mol, and the calculations were iterated until convergence was reached. The program GOLD 2.2 [49] was used for docking experiments, and all ligands, including both tautomers of the azo compounds, were docked. The standard default settings of GOLD were used, and the active site radius was set to 15 Å centered on the H_ζ atom of F79 of the avidin structure (PDB code 1VYO) and F81 of the AVR4 structure (PDB code 1Y55) [18]. The docking studies were performed based on the assumption that avidin binds the azo compounds in a manner similar to what has been observed for the avidin-HABA complex [20, 50]. In the docking studies, two distance constraints were used; otherwise, the program GOLD would not have managed to dock the azo compounds into the assumed conformation. The distance constraints were defined so that one of the carboxylate oxygens of the azo compounds was forced to be in the vicinity of the H_γ atom of S16 and the H_η atom of Y33 (both in avidin and AVR4). The range of separation was from 1.5 Å to 3.5 Å, with a spring constant of 5.0.

Visualization

The Bodil modeling environment [51] was used for the visualization of the crystal structures as well as for the visualization of the docking results. Figures were produced with PyMOL version 0.99 [52], and labels were added by Gimp 2.2.

Supplemental Data

Supplemental Data include additional Experimental Procedures, which contain data on the isolation and characterization of the synthesized azo compounds, and are available at http://www.chembiol.com/cgi/content/full/13/10/1029/DC1/.

Acknowledgments

We thank the staff of beamline X13 at the European Molecular Biology Laboratory (EMBL)/DESY (German Synchrotron Research Centre) Hamburg for excellent support. This project was supported by the European Community–Access to Research Infrastructure Action of the Improving Human Potential Programme to the EMBL Hamburg Outstation, contract no. HPRI-CT-1999-00017. We thank the CSC–Scientific Computing Ltd., Espoo, Finland for providing us with access to the program GOLD. This work was supported by grants from the Academy of Finland, the Technology Development Center (TEKES) of Finland, Sigrid Jusélius Foundation, National Graduate School in Informational and Structural Biology (ISB), and the Foundation of Åbo Akademi University (Center of Excellence Program in Cell Stress).

Received: June 2, 2006
Revised: August 11, 2006
Accepted: August 14, 2006
Published: October 20, 2006

References

1. Green, N.M. (1975). Avidin. Adv. Protein Chem. *29*, 85–133.
2. Green, N.M. (1990). Avidin and streptavidin. Methods Enzymol. *184*, 51–67.
3. Chilkoti, A., and Stayton, P.S. (1995). Molecular origins of the slow streptavidin-biotin dissociation kinetics. J. Am. Chem. Soc. *117*, 10622–10628.
4. Freitag, S., Chu, V., Penzotti, J., Klumb, L., To, R., Hyre, D., Trong, I., Lybrand, T., Stenkamp, R., and Stayton, P. (1999). A structural snapshot of an intermediate on the streptavidin-biotin dissociation pathway. Proc. Natl. Acad. Sci. USA *96*, 8384–8389.
5. Hyre, D.E., Le Trong, I., Freitag, S., Stenkamp, R.E., and Stayton, P.S. (2000). Ser45 plays an important role in managing both the equilibrium and transition state energetics of the streptavidin-biotin system. Protein Sci. *9*, 878–885.
6. Marttila, A., Airenne, K., Laitinen, O., Kulik, T., Bayer, E., Wilchek, M., and Kulomaa, M. (1998). Engineering of chicken avidin: a progressive series of reduced charge mutants. FEBS Lett. *441*, 313–317.
7. Nordlund, H.R., Laitinen, O.H., Uotila, S.T., Nyholm, T., Hytönen, V.P., Slotte, J.P., and Kulomaa, M.S. (2003). Enhancing the thermal stability of avidin. Introduction of disulfide bridges between subunit interfaces. J. Biol. Chem. *278*, 2479–2483.
8. Nordlund, H.R., Hytönen, V.P., Laitinen, O.H., Uotila, S.T., Niskanen, E.A., Savolainen, J., Porkka, E., and Kulomaa, M.S. (2003). Introduction of histidine residues into avidin subunit interfaces allows pH-dependent regulation of quaternary structure and biotin binding. FEBS Lett. *555*, 449–454.
9. Weber, P.C., Ohlendorf, D.H., Wendoloski, J.J., and Salemme, F.R. (1989). Structural origins of high-affinity biotin binding to streptavidin. Science *243*, 85–88.
10. Livnah, O., Bayer, E.A., Wilchek, M., and Sussman, J.L. (1993). Three-dimensional structures of avidin and the avidin-biotin complex. Proc. Natl. Acad. Sci. USA *90*, 5076–5080.
11. Hyre, D.E., Amon, L.M., Penzotti, J.E., Le Trong, I., Stenkamp, R.E., Lybrand, T.P., and Stayton, P.S. (2002). Early mechanistic events in biotin dissociation from streptavidin. Nat. Struct. Biol. *9*, 582–585.
12. Ahlroth, M.K., Kola, E.H., Ewald, D., Masabanda, J., Sazanov, A., Fries, R., and Kulomaa, M.S. (2000). Characterization and chromosomal localization of the chicken avidin gene family. Anim. Genet. *31*, 367–375.
13. Niskanen, E.A., Hytönen, V.P., Grapputo, A., Nordlund, H.R., Kulomaa, M.S., and Laitinen, O.H. (2005). Chicken genome analysis reveals novel genes encoding biotin-binding proteins related to avidin family. BMC Genomics *6*, 41.
14. Hytönen, V.P., Laitinen, O.H., Airenne, T.T., Kidron, H., Meltola, N.J., Porkka, E., Hörhä, J., Paldanius, T., Määttä, J.A., Nordlund, H.R., et al. (2004). Efficient production of active chicken avidin using a bacterial signal peptide in *Escherichia coli*. Biochem. J. *384*, 385–390.
15. Laitinen, O.H., Hytönen, V.P., Ahlroth, M.K., Pentikäinen, O.T., Gallagher, C., Nordlund, H.R., Ovod, V., Marttila, A.T., Porkka, E., Heino, S., et al. (2002). Chicken avidin-related proteins show altered biotin-binding and physico-chemical properties as compared with avidin. Biochem. J. *363*, 609–617.
16. Hytönen, V.P., Nyholm, T.K., Pentikainen, O.T., Vaarno, J., Porkka, E.J., Nordlund, H.R., Johnson, M.S., Slotte, J.P., Laitinen, O.H., and Kulomaa, M.S. (2004). Chicken avidin-related protein 4/5 shows superior thermal stability when compared with avidin while retaining high affinity to biotin. J. Biol. Chem. *279*, 9337–9343.
17. Hytönen, V.P., Maatta, J.A., Kidron, H., Halling, K.K., Horha, J., Kulomaa, T., Nyholm, T.K., Johnson, M.S., Salminen, T.A., Kulomaa, M.S., et al. (2005). Avidin related protein 2 shows unique structural and functional features among the avidin protein family. BMC Biotechnol. *5*, 28.
18. Eisenberg-Domovich, Y., Hytönen, V.P., Wilchek, M., Bayer, E.A., Kulomaa, M.S., and Livnah, O. (2005). High-resolution crystal structure of an avidin-related protein: insight into high-affinity biotin binding and protein stability. Acta Crystallogr. D Biol. Crystallogr. *61*, 528–538.
19. Green, N.M. (1970). Spectrophotometric determination of avidin and streptavidin. Methods Enzymol. *18*, 418–424.
20. Livnah, O., Bayer, A., Wilchek, M., and Sussman, J.L. (1993). The structure of the complex between avidin and the dye, 2-(4′-hydroxyazobenzene) benzoic acid (HABA). FEBS Lett. *328*, 165–168.
21. Weber, P.C., Wendoloski, J.J., Pantoliano, M.W., and Salemme, F.R. (1992). Crystallogaphic and thermodynamic comparison of natural and synthetic ligands bound to streptavidin. J. Am. Chem. Soc. *114*, 3197–3200.
22. Gonzales, M., Argarana, C.E., and Fidelio, G.D. (1999). Extremely high thermal stability of streptavidin and avidin upon biotin binding. Biomol. Eng. *16*, 67–72.
23. Williams, D.H., Stephens, E., and Zhou, M. (2003). Ligand binding energy and catalytic efficiency from improved packing within receptors and enzymes. J. Mol. Biol. *329*, 389–399.
24. Pugliese, L., Malcovati, M., Coda, A., and Bolognesi, M. (1994). Crystal structure of apo-avidin from hen egg-white. J. Mol. Biol. *235*, 42–46.
25. Korndörfer, I.P., and Skerra, A. (2002). Improved affinity of engineered streptavidin for the Strep-tag II peptide is due to a fixed open conformation of the lid-like loop at the binding site. Protein Sci. *11*, 883–893.
26. Weber, P.C., Pantoliano, M.W., Simons, D.M., and Salemme, F.R. (1994). Structure-based design of synthetic azobenzene ligands for streptavidin. J. Am. Chem. Soc. *116*, 2717–2724.
27. Huberman, T., Eisenberg-Domovich, Y., Gitlin, G., Kulik, T., Bayer, E.A., Wilchek, M., and Livnah, O. (2001). Chicken avidin exhibits pseudo-catalytic properties. Biochemical, structural, and electrostatic consequences. J. Biol. Chem. *276*, 32031–32039.
28. Nardone, E., Rosano, C., Santambrogio, P., Curnis, F., Corti, A., Magni, F., Siccardi, A.G., Paganelli, G., Losso, R., Apreda, B., et al. (1998). Biochemical characterization and crystal structure of a recombinant hen avidin and its acidic mutant expressed in *Escherichia coli*. Eur. J. Biochem. *256*, 453–460.
29. Pazy, Y., Eisenberg-Domovich, Y., Laitinen, O.H., Kulomaa, M.S., Bayer, E.A., Wilchek, M., and Livnah, O. (2003). Dimer-tetramer transition between solution and crystalline states of streptavidin and avidin mutants. J. Bacteriol. *185*, 4050–4056.
30. Pazy, Y., Kulik, T., Bayer, E.A., Wilchek, M., and Livnah, O. (2002). Ligand exchange between proteins. Exchange of biotin and biotin derivatives between avidin and streptavidin. J. Biol. Chem. *277*, 30892–30900.
31. Pugliese, L., Coda, A., Malcovati, M., and Bolognesi, M. (1993). Three-dimensional structure of the tetragonal crystal form of egg-white avidin in its functional complex with biotin at 2.7 Å resolution. J. Mol. Biol. *231*, 698–710.

32. Brandts, J.F., and Lin, L.N. (1990). Study of strong to ultratight protein interactions using differential scanning calorimetry. Biochemistry *29*, 6927–6940.
33. Green, N.M. (1965). A spectrophotometric assay for avidin and biotin based on binding of dyes by avidin. Biochem. J. *94*, 23C–24C.
34. Mock, D.M., Lankford, G., and Horowitz, P. (1988). A study of the interaction of avidin with 2-anilinonaphtalene-6-sulfonic acid as a probe of the biotin binding site. Biochim. Biophys. Acta *956*, 23–29.
35. Pazy, Y., Raboy, B., Matto, M., Bayer, E.A., Wilchek, M., and Livnah, O. (2003). Structure-based rational design of streptavidin mutants with pseudo-catalytic activity. J. Biol. Chem. *278*, 7131–7134.
36. Ellison, D., Hinton, J., Hubbard, S.J., and Beynon, R.J. (1995). Limited proteolysis of native proteins: the interaction between avidin and proteinase K. Protein Sci. *4*, 1337–1345.
37. Hytönen, V.P., Määttä, J.A., Nyholm, T.K., Livnah, O., Eisenberg-Domovich, Y., Hyre, D., Nordlund, H.R., Hörhä, J., Niskanen, E.A., Paldanius, T., et al. (2005). Design and construction of highly stable, protease-resistant chimeric avidins. J. Biol. Chem. *280*, 10228–10233.
38. Kabsch, W. (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. J. Appl. Crystallogr. *26*, 795–800.
39. CCP4 (Collaborative Computational Project, Number 4) (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. *50*, 760–763.
40. Navaza, J. (1994). AMoRe—an automated package for molecular replacement. Acta Crystallogr. A *50*, 157–163.
41. Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D Biol. Crystallogr. *53*, 240–255.
42. Howlin, B., Butler, S.A., Moss, D.S., Harris, G.W., and Driessen, H.P.C. (1993). TlsanI—Tls parameter-analysis program for segmented anisotropic refinement of macromolecular structures. J. Appl. Crystallogr. *26*, 622–624.
43. Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A *47*, 110–119.
44. Lamzin, V.S., and Wilson, K.S. (1993). Automated refinement of protein models. Acta Crystallogr. D Biol. Crystallogr. *49*, 129–147.
45. Laskowski, R.A., Macarthur, M.W., Moss, D.S., and Thornton, J.M. (1993). Procheck—a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. *26*, 283–291.
46. Vriend, G. (1990). WHAT IF: a molecular modeling and drug design program. J. Mol. Graph. *8*, 52–56, 29.
47. Halgren, T.A. (1990). Maximally diagonal force constants in dependent angle-bending coordinates. Implications for the design of empirical force fields. J. Am. Chem. Soc. *112*, 4710–4723.
48. Halgren, T.A. (1999). MMFF VI. MMFF94s option for energy minimization studies. J. Comput. Chem. *20*, 720–729.
49. Jones, G., Willett, P., Glen, R.C., Leach, A.R., and Taylor, R. (1997). Development and validation of a genetic algorithm for flexible docking. J. Mol. Biol. *267*, 727–748.
50. Kuhn, B., and Kollman, P.A. (2000). Binding of a diverse set of ligands to avidin and streptavidin: an accurate quantitative prediction of their relative affinities by a combination of molecular mechanics and continuum solvent models. J. Med. Chem. *43*, 3786–3791.
51. Lehtonen, J.V., Still, D.-J., Rantanen, V.-V., Ekholm, J., Björklund, D., Iftikhar, Z., Huhtala, M., Repo, S., Jussila, A., Jaakkola, J., et al. (2004). BODIL: a molecular modeling environment for structure-function analysis and drug design. J. Comput. Aided Mol. Des. *18*, 401–419.
52. DeLano, W.L. (2002). The PyMOL Molecular Graphics System (http://www.pymol.org).
53. Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2000). The Protein Data Bank. Nucleic Acids Res. *28*, 235–242.
54. Barton, G.J. (1993). ALSCRIPT a tool to format multiple sequence alignments. Protein Eng. *6*, 37–40.

Accession Numbers

Coordinates have been deposited in the Protein Data Bank [53] with accession code 1VYO.

OPEN ACCESS Freely available online

PLoS one

Construction of Chimeric Dual-Chain Avidin by Tandem Fusion of the Related Avidins

Tiina A. Riihimäki[1], Sampo Kukkurainen[1☯], Suvi Varjonen[1☯¤a], Jarno Hörhä[2¤b], Thomas K. M. Nyholm[3], Markku S. Kulomaa[1,2], Vesa P. Hytönen[1,2]*

1 Institute of Biomedical Technology, University of Tampere and Tampere University Hospital, Tampere, Finland, **2** Department of Biological and Environmental Science, University of Jyväskylä, Jyväskylä, Finland, **3** Department of Biochemistry and Pharmacy, Åbo Akademi University, Turku, Finland

Abstract

Background: Avidin is a chicken egg-white protein with high affinity to vitamin H, also known as D-biotin. Many applications in life science research are based on this strong interaction. Avidin is a homotetrameric protein, which promotes its modification to symmetrical entities. Dual-chain avidin, a genetically engineered avidin form, has two circularly permuted chicken avidin monomers that are tandem-fused into one polypeptide chain. This form of avidin enables independent modification of the two domains, including the two biotin-binding pockets; however, decreased yields in protein production, compared to wt avidin, and complicated genetic manipulation of two highly similar DNA sequences in the tandem gene have limited the use of dual-chain avidin in biotechnological applications.

Principal Findings: To overcome challenges associated with the original dual-chain avidin, we developed chimeric dual-chain avidin, which is a tandem fusion of avidin and avidin-related protein 4 (AVR4), another member of the chicken avidin gene family. We observed an increase in protein production and better thermal stability, compared with the original dual-chain avidin. Additionally, PCR amplification of the hybrid gene was more efficient, thus enabling more convenient and straightforward modification of the dual-chain avidin. When studied closer, the generated chimeric dual-chain avidin showed biphasic biotin dissociation.

Significance: The improved dual-chain avidin introduced here increases its potential for future applications. This molecule offers a valuable base for developing bi-functional avidin tools for bioseparation, carrier proteins, and nanoscale adapters. Additionally, this strategy could be helpful when generating hetero-oligomers from other oligomeric proteins with high structural similarity.

Citation: Riihimäki TA, Kukkurainen S, Varjonen S, Hörhä J, Nyholm TKM, et al. (2011) Construction of Chimeric Dual-Chain Avidin by Tandem Fusion of the Related Avidins. PLoS ONE 6(5): e20535. doi:10.1371/journal.pone.0020535

Editor: Narcis Fernandez-Fuentes, Leeds Institute of Molecular Medicine, United Kingdom

Received March 21, 2011; **Accepted** May 3, 2011; **Published** May 31, 2011

Funding: Funding by the Academy of Finland (http://www.aka.fi/en-GB/A/; project numbers 115976 and 121236), National Doctoral Programme in Informational and Structural Biology (ISB, http://web.abo.fi/isb/) and Tampere Graduate Program in Biomedicine and Biotechnology (TGPBB, http://www.uta.fi/tutkijakoulut/tgsbb/) is greatly appreciated. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: vesa.hytonen@uta.fi

☯ These authors contributed equally to this work.

¤a Current address: VTT Technical Research Centre of Finland, Espoo, Finland
¤b Current address: Institute of Biomedicine/Physiology, University of Helsinki, Helsinki, Finland

Introduction

Improving the performance and accuracy of molecular tools used in life science research is essential for developing better and more precise methods. Chicken avidin (AVD) and its bacterial analogue streptavidin (SA) from *Streptomyces avidinii*, collectively called (strept)avidin, are proteins widely used in life science research applications. (Strept)avidin has been used for quantitative measurements by radioligand-binding methods [1], enzyme assays [2,3] and photometric/fluorometric methods [4–6]. AVD has also been successfully used in biosensors as an immobilization platform [7]. The specific characteristics of AVD, such as the high positive charge (pI 10.5) and high biotin-binding affinity ($K_d \approx 10^{-15}$ M), have resulted in a number of different drug-targeting applications that use the (strept)avidin-biotin interaction [8].

Although widely used, (strept)avidin's homotetrameric structure restricts its use in some applications. To overcome this limitation, dual-chain avidin (dcAVD) was generated [9] by fusing two circularly permuted chicken avidin monomers into one polypeptide chain. This molecular engineering approach produced a protein that resembles wt avidin in 3-D structure [10] but allows independent manipulation of ligand-binding sites within a single protein particle. To create dual-affinity derivatives, dcAVD was modified by site-directed mutagenesis. This form of dcAVD exhibited a tight biotin affinity for two binding sites, whereas the other two binding sites had reduced affinity due to the mutations [11]. Recently, a point mutation S16C was targeted into one of the biotin-binding sites of dcAVD [12]. The introduction of a chemically active thiol group into the binding site made it possible to selectively control the biotin-

binding activity of the dcAVD domains by a mild chemical treatment.

Despite recent improvements, some challenges that could limit the use of dcAVD persist. For instance, PCR amplification of the dcAVD-encoding sequence has been challenging because primers recognize complementary sequences from both subunits, which would produce several different PCR products. This same issue also limits targeted mutagenesis of dcAVD. This study uses a new approach for developing the dcAVDs. Instead of constructing a tandem gene by combining two differently circularly permuted chicken avidins [9] or streptavidins [13], two related avidin genes were used as building blocks for the chimeric tandem fusion. The raw materials of the chimeras were selected from a group of homologous proteins. Streptavidin (SA), the most widely used protein in biotechnological applications in the avidin family, and avidin-related proteins 2 and 4 (AVR2 and AVR4) [14,15] were selected as fusion partners for the circularly permuted chicken avidin. By combining homologous genes into the chimeric tandem gene, we were able to address the problems associated with the original dcAVD. According to our knowledge, this is the first study showing such forced hybridization of different types of biotin-binding proteins into an oligomeric assembly.

Results and Discussion

The chimeric dcAVD fusions were designed according to previously described principles [9]. The loop connecting the original termini of the circularly permuted SA, AVR4, and AVR2 was shortened when compared to the original dcAVD, as presented in Figure 1. This shortening was performed because the loop was largely invisible in the X-ray analysis of dcAVD, indicating high mobility of the loop region [10].

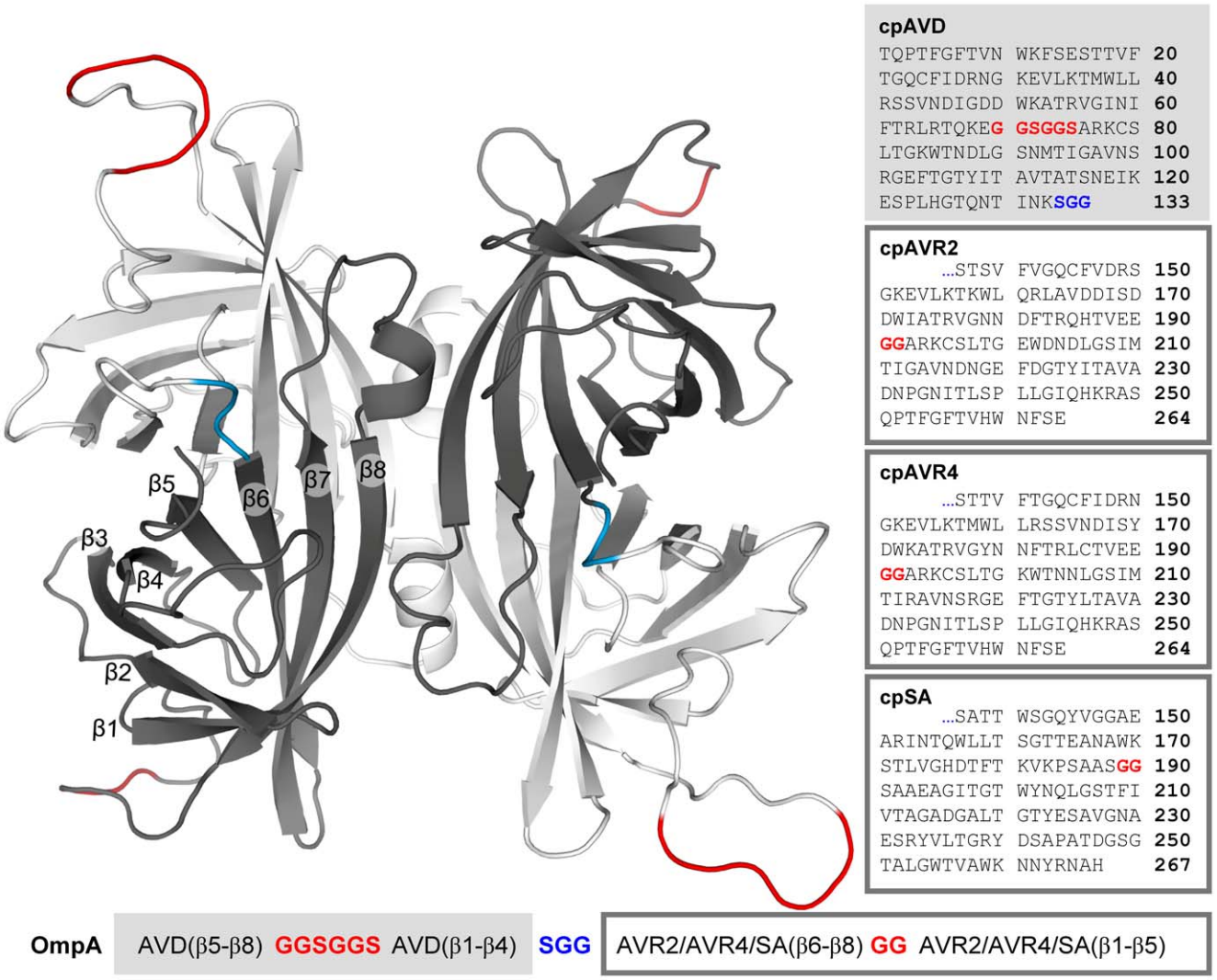


Figure 1. The homology model of dcAVD/AVR4 and the sequences of chimeric dcAVDs. The molecular model of dcAVD/AVR4 is generated by exploiting the existing 3-D structures of the dcAVD and AVR4. In the model, cpAVD is illustrated in light gray, and cpAVR4 is illustrated in dark gray. Amino acid sequence of cpAVD is in the light gray box, and the amino acid sequences of cpAVR2, cpAVR4, and cpSA are in the dark gray boxes. The linkers and the corresponding linker sequences connecting the original termini are shown in red. The linkers connecting the circularly permuted subunits and the corresponding linker sequences are shown in blue. The schematic representation of the protein expression cassette is in the bottom of the figure.
doi:10.1371/journal.pone.0020535.g001

Chimeric dcAVD genes performed well in PCR

Two conditions were used to evaluate performance of the dcAVD forms in PCR amplification. When primers recognizing the sequence that flank the tandem fusion gene were used, the amplification resulted in no notable differences between dcAVD and chimeric fusions. Amplification of the four dual-chain genes produced appropriately sized (~1000 bp) PCR products; however, a shorter product (700 bp) was clearly observed, except for dcAVD/SA (not shown). The short product could be a result of the amplification of a self-hybridized tandem gene. No clear differences were observed for reactions using the different thermal cycling parameters.

In contrast, when primers recognizing regions inside the tandem fusion gene were used, a clear difference in the behavior between dcAVD and the chimeric fusions was detected (Figure 2A). A number of side products were generated during the PCR amplification of the dcAVD gene, which was probably due to primers binding to multiple positions in the tandem gene and homologous recombination during the amplification process. In contrast, for the chimeric fusions, only two main PCR products were detected, which were the appropriately sized products. The dcAVD/SA showed the most efficient amplification. Using the chimeric fusion genes significantly improves PCR amplification and would allow targeting of mutagenesis to only a part of the gene, such as using the Quik-Change (Stratagene, La Jolla, CA, US) protocol. Moreover, this would enable the broader modification of dcAVD molecules, including targeted random mutagenesis of several amino acids.

The chimeric fusion dcAVD/AVR4 can be efficiently expressed in *E. coli*

All proteins selected as chimeric fusion partners resemble avidin in their fold and 3-D structure [16–18]. Therefore, we assumed that chimeric dual-chain avidins could be produced where one circularly permuted subunit would be based on the avidin sequence, and the other subunit would be based on another biotin-binding protein (AVR2, AVR4, SA). The chimeric dual-chain fusions were produced in *E. coli* using the periplasmic signal peptide from the *Bordetella avium* ompA protein [19]. The chimeric fusion dcAVD/AVR4 produced the best levels of the proteins analyzed, and the protein was almost entirely intact (~35 kDa) after 2-iminobiotin affinity chromatography (Figure 2B). To further study the usability of this chimeric protein, a pilot-scale expression of the dcAVD/AVR4 protein was performed in a 7.5 L fermentor. The pilot-scale fed-batch fermentation in standard LB medium yielded greater than 5 mg of pure dcAVD/AVR4 protein per liter of production medium with low amounts of protein fragments (Figure S1).

When dcAVD/AVR2 and dcAVD/SA proteins were produced, there was a significant amount of fragmented (~15 kDa) products in the SDS-PAGE analysis (Figure 2B). The low-sequence identity of SA with avidin (only ~30%, AVR's identity with avidin ~80% [20]) may explain the modest performance during the production of the chimeric dual-chain fusion with SA. Because some full-length protein was detected in the SDS-PAGE analysis of dcAVD/SA (Figure 2B), it might be possible to genetically tune this AVD/SA-hybrid to enhance its performance. The possible targets for such optimizations are discussed later in the text.

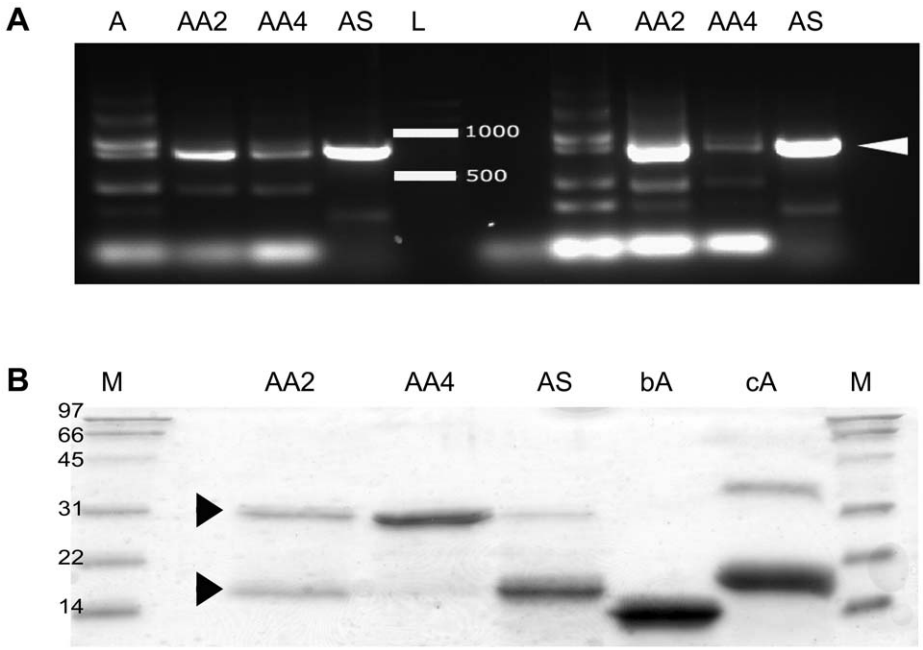


Figure 2. The performance of the chimeric dual-chain avidins in the PCR analysis and in *E. coli* expression. The usability of the generated chimeric dual chain avidin fusions was evaluated by amplifying the fusion genes by PCR and by expressing the proteins in *E. coli*. A) In the PCR analysis with primers recognizing regions inside the chimeric dcAVD genes (PCR 2), a clear difference between the behavior of the dcAVD genes (A) and the chimeric fusion genes (AA2, dcAVD/AVR2; AA4, dcAVD/AVR4; SA, dcAVD/SA) was detected. The chimeric fusion genes showed only two main PCR products; the appropriately sized product had the highest concentration. When the dcAVD gene was used as a template, several different-sized products were produced. The results from the PCR2 reaction (see Table S1) from two different conditions (I, II) are shown in the figure. (L, 1 kb DNA ladder). B) SDS-PAGE analysis of the purified chimeric dcAVDs showed that dcAVD/AVR4 was the most successfully expressed in its functional form in *E. coli*. The upper arrowhead indicates the location of the intact protein, and the lower arrowhead indicates the location of the proteolytically cleaved product. (M, molecular weight standard; AA2, dc-AVD-AVR2; AA4, dc-AVD-AVR4; AS, dc-AVD-SA, bA, chicken avidin control sample (protein expressed in *E. coli* [19]); cA, chicken avidin control sample).
doi:10.1371/journal.pone.0020535.g002

Size-exclusion chromatography was used to determine the oligomeric state of the produced proteins. The analysis revealed that the dcAVD/AVR4 protein was mostly in a pseudo-tetrameric form (Figure S2). Some higher molecular weight species were also detected, which could be explained by the oligomerization of pseudo-tetramers often detected in wt avidin samples [19]. Size-exclusion chromatography analysis of dcAVD/AVR2 revealed a significant proportion of clearly higher molecular weight species than dcAVD/AVR4. Interestingly, both dcAVD/SA and dcAVD/AVR2 appeared mostly in a pseudo-tetrameric form in gel-filtration analysis (results not shown). This result may indicate that the truncated protein forms observed in the SDS-PAGE analysis (Figure 2B) might be able to form oligomeric species, resulting in homotetrameric proteins. In any case, further studies are needed to better understand the properties of dcAVD/AVR2 and dcAVD/SA. In this study, dcAVD/AVR4 protein was selected to further biochemical analyses.

Differences in the subunit interfaces reveal possible reasons for the characteristics of chimeric dual-chain avidins

To analyze the possible reasons for the behavioral differences of the chimeric dual-chain avidins, the fusion proteins were modeled based on previously determined 3-D structures of dcAVD (PDB 2C4I), SA (PDB 1MK5), and AVRs (AVR2 (PDB 1WBI), AVR4 (PDB 1Y53)). The model of dcAVD/AVR4 is presented in Figure 1. Molecular dynamics (MD) were performed for the predicted chimeric dual-chain avidin models. The simulations were performed in explicit water using the CHARMM force field. The interaction energy during the MD simulation was measured between subunits, which is the most obvious region in the structure that would cause problems in the dcAVD assembly. In the MD simulation analyses, dcAVD/SA had clearly the least favorable electrostatic interaction energy (Figure 3A), whereas there were no significant differences in the van der Waals energies between chimeric dcAVD forms (Figure 3B). A closer inspection of the MD simulation data of dcAVD/SA revealed three putative electrostatically repulsive interactions (Figures 3C and 3D). These residues (D247-E15-D241, K12-R239 and D27-E165) are potential targets for further engineering of dcAVD/SA to improve its characteristics.

To investigate how the chimeric dcAVD forms differ in terms of loop dynamics in MD simulations, we analyzed a root mean square fluctuation (RMSF), measuring main chain motion for a 10 ps time window (Figure 4). The analysis revealed that the linkers connecting the original termini of the domains (Figure 1, red linkers) were highly mobile, particularly in the avidin domain (GGSGGS, residues 70–75 connecting the original termini). We also detected regions behaving differently between the chimeric dcAVD forms; for dcAVD/SA, there were unique mobile regions in the loop between the strands β7 and β8 (residues 164–169) and in the loop between the strands β4 and β5 (residues 243–249) that corresponded to the potential electrostatic repulsion (Figure 3C and 3D). Overall, however, no dramatic differences were observed in the loop mobility between different dcAVD forms, suggesting that there were no significant problems in the molecular design of the chimeric dcAVDs.

The chimeric dcAVD/AVR4 showed an increase in thermal stability when compared to dcAVD

A thermal stability assessment of the dcAVD/AVR4 by SDS-PAGE revealed a significant improvement when compared to that of dcAVD. The determined transition temperature of subunit dissociation (T_r) was 25°C higher for dcAVD/AVR4 (Table 1). The presence of D-biotin in the binding site stabilized the dcAVD/AVR4 quaternary structure, and the determined transition temperature of oligomeric disassembly (T_r) increased from 65°C to 85°C; however, when compared to dcAVD, the increase in transition temperature due to ligand binding was less (Table 1).

Differential scanning calorimetry (DSC) was used to analyse the thermodynamics of the heat-induced unfolding. In a DSC analysis of dcAVD/AVR4, a two-phase melting profile was observed both with and without biotin (Figure 5B). This melting profile was not detected in the dcAVD samples, in which the heat induced unfolding resulted in a single peak in the thermogram (Figure 5A). The thermograms recorded with dcAVD/AVR4 samples were deconvoluted in order to separate the two peaks (Figure 5C & Figure 5D). The main peak in the dcAVD/AVR4 thermogram revealed a melting point (T_m) at 91.3°C (Figure 5C), which is about 11°C higher compared to that measured for dcAVD (80.2°C). The smaller secondary peak showed a melting point at 86.3°C. In the presence of biotin, the main peak showed a melting point at 112.3°C, and a secondary peak at 107.8°C (Figure 5D). Interestingly, in the presence of biotin, dcAVD/AVR4 was denatured at a lower temperature compared to the melting temperature (115.9°C) of dcAVD. Overall, dcAVD/AVR4 had improved thermal stability in the absence of biotin, but the addition of biotin did not produce as significant thermal stabilization as in the case of dcAVD or wt AVD. Therefore, the exchange of the cp65-subunit of dcAVD with a circularly permuted AVR4 subunit increased the thermal stability of the chimeric fusion protein, while the other domain of the tandem gene remained unchanged (AVD-derived cp54). The lower biotin-binding affinity of the introduced AVR4-derived subunit reflected the thermal stability of the whole fusion protein in the presence of biotin. These results are clear indications of structural cooperativity between the subunits of the tetramer during thermal unfolding; however, previous studies have shown that (strept)avidin has relatively little or no structural cooperativity between subunits [21–23]. One reason for the apparently low cooperativity has been attributed to the subunit exchange between partially unfolded proteins [24]. For dcAVD, the covalent attachment of the subunits might block or at least significantly reduce the subunit exchange in the thermal unfolding process. Therefore, the dual chain concept allows a novel type of approach in studies elucidating the unfolding mechanisms of avidin proteins. The unfolding process was irreversible, which is typical for avidin proteins.

The biotin-binding properties of dcAVD/AVR4 differed from those of the parental proteins

Fluorescently labeled biotin and surface plasmon resonance (SPR) biosensor were used to study the biotin-binding characteristics of dcAVD/AVR4. In the experiment with fluorescently labeled biotin a bi-phasic ligand-dissociation process was detected, where roughly 50% of the protein subunits released biotin with rapid ($k_{diss} = 1.1 \times 10^{-3}\ s^{-1}$) dissociation kinetics (Figure 6). It is probable that the rapid biotin-dissociation phase was associated with the AVR4-derived cp65 domain. As we have previously shown, the biotin-binding affinity of the circularly permuted avidin cp65 is slightly less than wt AVD. In contrast, the cp54 version appears to behave more like wt AVD in terms of biotin-binding [9]. The previous studies have shown that the AVR4 protein has a slightly lower biotin-binding affinity ($K_d = 3.6 \times 10^{-14}$ M) than avidin ($K_d = 1.1 \times 10^{-16}$ M) [25]; however, the measured increase in the biotin dissociation rate of dcAVD/AVR4 was higher than expected based on previous studies, possibly reflecting the

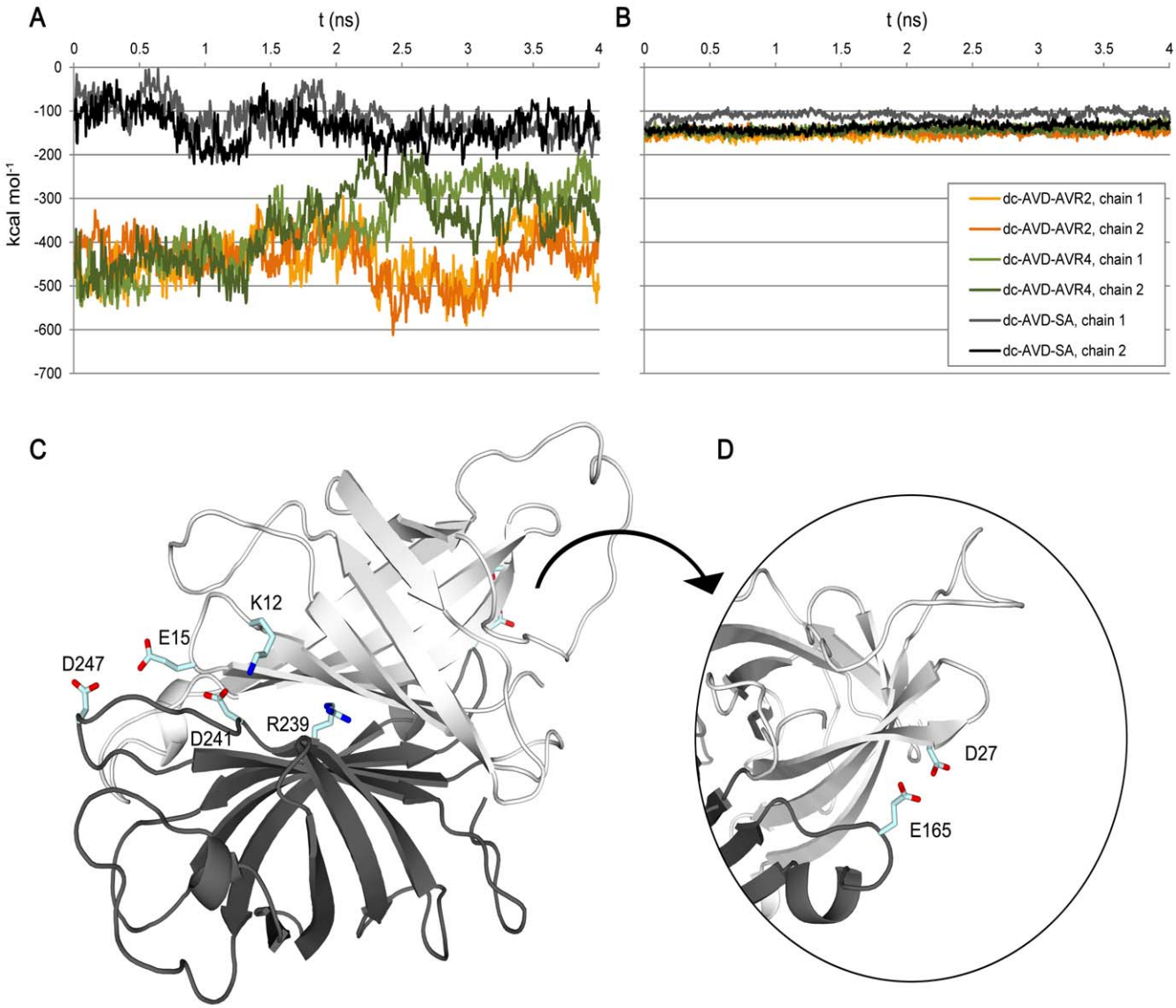


Figure 3. Interactions between the subunit interfaces of chimeric dcAVD fusions by MD simulation. The electrostatic interaction energy was measured between the circularly permuted subunits in chimeric dcAVDs (A). The analysis was performed for both subunit pairs independently, and the interaction energy is plotted over the 4-ns MD simulation. The van der Waals energy between subunits was measured during the simulation time (B). The potential sources of electrostatic repulsion for dcAVD/SA were examined by visual inspection of the MD simulation data. Three clusters of residues (D247-E15-D241, K12-R239 and D27-E165) potentially causing electrostatic repulsion between cpAVD and cpSA were detected. These residues are shown in a liquorice representation (C, D). The figures were prepared using the program PyMOL (www.pymol.org) and numbered according to Figure 1. All the calculations were performed with 5-ps resolution.
doi:10.1371/journal.pone.0020535.g003

cooperativity between different parts of the biotin-binding site [26,27].

In the SPR analysis using the Biacore X instrument, the biotin derivative 2–iminobiotin was coupled to the sensor chip using amino-coupling chemistry. For the dcAVD/AVR4, we measured an apparent equilibrium dissociation constant of 9.6×10^{-6} M, which is greater than the constant measured for the parental proteins (wtAVD $K_d = 1.2\times10^{-7}$ M, AVR4(C112S) $K_d = 7.2\times10^{-7}$ M). The binding of dcAVD strongly resembled the wtAVD that was analyzed in the SPR assay in our preceding study [9].

The results from the fluorescent biotin interaction and the SPR analyses were consistent with each other; however, for the SPR measurements, the 2-iminobiotin ligand was attached onto the surface, and therefore, the interaction between the immobilized ligand and free protein was measured. For the measurements using the fluorescently labeled biotin, the ligand moved freely in solution. Therefore, SPR analysis could overestimate the binding affinity of chimeric dcAVDs because the protein might preferably bind the immobilized ligand with the subunit that has the greater binding affinity.

The biotin-binding properties of dcAVD/AVR4 and those measured for dcAVDs carrying point mutations in the biotin-binding site [11] suggested that the modification of the loop connecting beta strands 5 and 6 in the C-terminal domain of dual-chain avidin had negative effects on biotin-binding. In both studies, the dissociation rate constant for the C-terminal domain (cp 65, see Figure 1) was decreased more than expected, based on

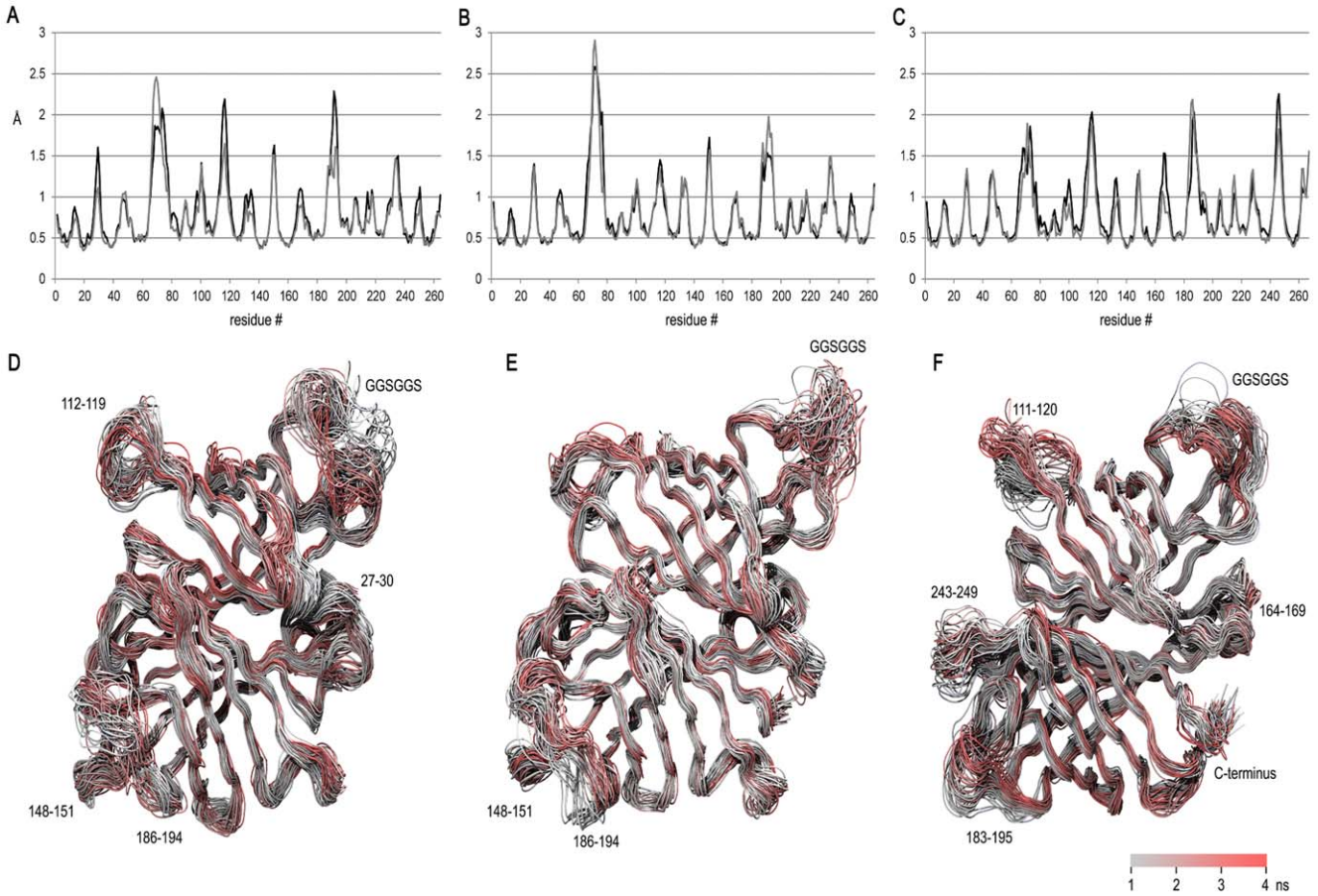


Figure 4. Local dynamics in the chimeric dcAVD fusions measured by MD simulation. To probe the local structural dynamics, the root mean square fluctuation (RMSF) per residue was measured for a short time window (10 ps) for the last 3 ns of the 4-ns MD simulation. The resulting values were averaged and plotted in graphs A–C (A: dcAVD/AVR2; B: dcAVD/AVR4; C: dcAVD/SA). The dynamics of the structure are illustrated by plotting 50 superimposed structural snapshots along the 4-ns simulation (D–F). The loops showing a high amount of structural fluctuation (RMSF>1.5 Å) are indicated by numbers referring to the amino acid sequence (see also Figure 1). The structural snapshots are colored according to timestep, as illustrated by the scale bar (please note that the color scale is illustrative only because of the rendering method). Figures D–F were prepared using the program VMD 1.8.7 [35].
doi:10.1371/journal.pone.0020535.g004

the sequence of the domain. Therefore, it would be logical to engineer the C-terminal domain of the dcAVD for novel characteristics while preserving the wt AVD-like characteristics of the N-terminal domain (cp54). This would exploit the existing biotinylated molecular tools maximally in dcAVD-based applications.

Table 1. Transition temperatures of the subunit dissociation (T_r) and thermal unfolding (T_m) determined by SDS-PAGE and DSC.

	SDS-PAGE		DSC	
	T_r (°C)	$T_r^{(+BTN)}$ (°C)	T_m (°C)	$T_m^{(+BTN)}$ (°C)
AVD	60[a]	95[a]	84[b]	117[b]
AVR4(C112S)	90	95	110[b]	127[b]
dcAVD	40	75[c]	80[a]	116[a]
dcAVD/AVR4	70	85	*86*/91[d]	*108*/112[d]

[a]Value from [19];
[b][25];
[c][9];
[d](two-peak analysis was applied to the data, and the minor peak is shown in *italics*).
doi:10.1371/journal.pone.0020535.t001

Conclusions

Dual-chain avidin is an example of the complicated genetic engineering possible with the robust structure of avidin [9]. In the current study, the usability of dcAVD was improved by creating chimeric dual-chain avidin proteins. As a fusion partner with wt AVD, we used three different proteins from the avidin protein family: streptavidin, AVR2 and AVR4. The most successful was a chimeric tandem fusion of circularly permutated wt AVD and circularly permuted AVR4. Enhanced protein expression and thermal stability resulted when compared to the original dcAVD. Also, the PCR amplification was more straightforward when using chimeric dual chain fusion. Closer analyses of dcAVD/AVR4 protein showed that the molecule exhibited heterogenous biotin-binding. This might be advantageous in applications where two different biotin-binding affinities are needed. To further this technique, the dcAVD/AVR4 format can be combined with other genetically engineered avidins with modified or completely

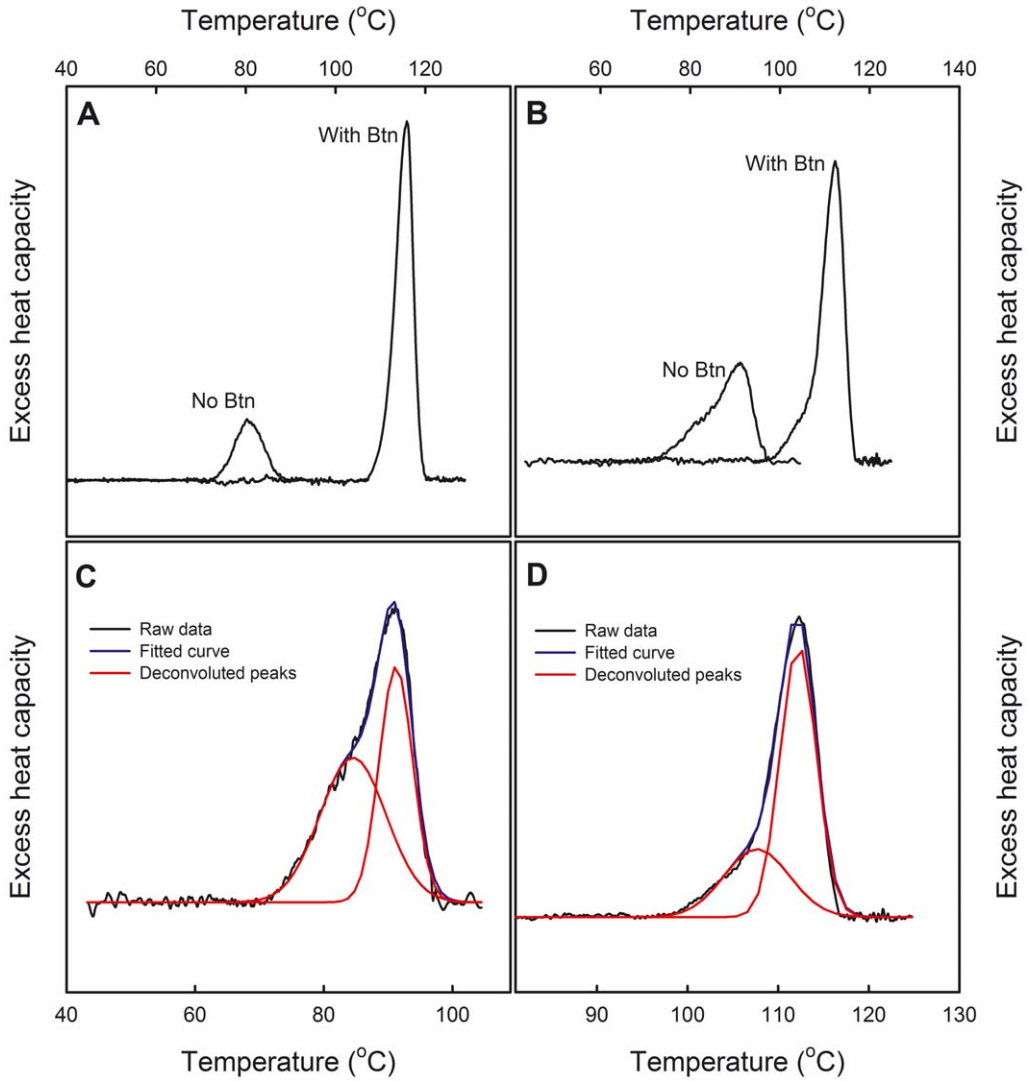


Figure 5. DSC analysis shows the biphasic thermal denaturation mode of the dcAvd/AVR4 protein. Heat-induced unfolding of the proteins was analysed by differential scanning calorimetry. The measured thermogram of dcAVD (A) and dcAVD/AVR4 (B) is shown without and with biotin (With Btn). Deconvoluted thermograms of dcAVD/AVR4 without (C) and with (D) biotin are also shown. The thermogram of dcAVD/AVR4 shows a melting point (T_m) at 91.3°C (C), which is about 11°C greater than for dcAVD (80.2°C, [11]). The smaller secondary peak shows a melting point at 86.3°C. Biphasic thermal denaturation mode is also detected in the presence of biotin (D); the melting point of the main peak is at 112.3°C, and the secondary peak is at 107.8°C. Interestingly, in the presence of biotin, dcAVD/AVR4 was denatured at a slightly lower temperature compared with dcAVD 115.9°C [11].
doi:10.1371/journal.pone.0020535.g005

new binding properties. For instance, avidins with steroid-binding have been developed in our group by random mutagenesis and selected by phage display (Riihimäki et al., manuscript). By combining these modified avidins to a dcAVD or single-chain (scAVD, [28]) avidin platform, it will be possible to develop avidin-based receptors with alternative binding affinities or with multiple ligand specificities to be used in *in vitro* diagnostics or in nanotechnology.

Materials and Methods

Evaluation of the PCR performance of the chimeric dcAvd fusions

PCR amplification of the chimeric tandem fusion genes was performed to evaluate their suitability for genetic modification. The first PCR experiment was performed using four conditions; the second PCR experiment was performed using two conditions (Table S1). *Pfu* polymerase (Fermentas) was used and reactions were performed following the manufacturer's instructions. All PCR products were analyzed by 1.5% agarose gel electrophoresis using a 1-kbp DNA ladder as a standard.

Production of chimeric dcAVD fusion proteins

DNA constructs of circularly permutated avidin cp54 were fused to circularly permuted streptavidin, AVR2 and AVR4 (cp65 SA/AVR2/AVR4) [9]. Circularly permuted genes were inserted into the pBVboostFG plasmid containing the region encoding the OmpA signal sequence, and the plasmid was transformed into chemically competent *E. coli* TOP10 cells (Invitrogen) by the standard heat-shock method. The plasmids were harvested using the Qiagen Plasmid Miniprep kit according to the manufacturer's instructions, and the inserts were verified by DNA sequencing.

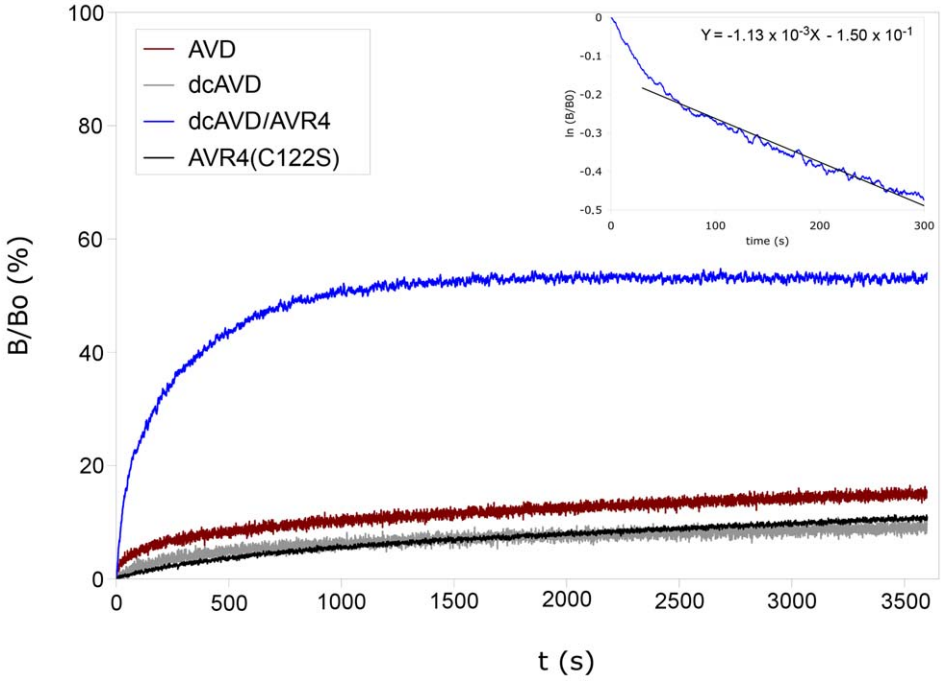


Figure 6. Determination of ligand dissociation kinetics with a fluorescent-biotin conjugate. The dissociation of fluorescently labeled biotin from various avidin forms was studied by replacing the labeled biotin with an excess of free biotin. AVD, AVR4(C122S) and dcAVD showed a slow dissociation. For dcAVD/AVR4, a clearly biphasic dissociation process was observed. The analysis of the first part of the measurement (0–300 s, inset) reveals an estimate for the dissociation rate constant of 1.13×10^{-3} s^{-1} for the rapid dissociation phase of dcAVD/AVR4, which is about 100× greater when compared to avidin or AVR4(C122S).
doi:10.1371/journal.pone.0020535.g006

The amino acid sequences of the recombinant proteins and the cloning cassette are shown in Figure 1.

For protein production, the plasmids were transformed into *E. coli* BL21-AI cells (Invitrogen) by the heat-shock method. After overnight induction at 26°C, the avidin proteins were purified by affinity chromatography on a 2-iminobiotin column (Affiland, Liège, Belgium) [19,29]. The avidin concentration was measured with a NanoDrop™ 3300 spectrometer using a molar absorption coefficient (for dcAVD/AVR4; $\varepsilon = 50\ 880\ M^{-1}\ cm^{-1}$ and M = 29 045 g/mol, for dcAVD/AVR2; $\varepsilon = 48\ 320\ M^{-1}\ cm^{-1}$ and M = 28 866 g/mol and for dcAVD/SA $\varepsilon = 65\ 980\ M^{-1}\ cm^{-1}$ and M = 28 316 g/mol) at 280 nm. The produced hybrid dcAVD proteins were analyzed by SDS-PAGE stained with Coomassie Brilliant Blue.

Production of dcAvd/AVR4 by pilot-scale fermentation

For pilot-scale fermentation of dcAVD/AVR4 protein, the Infors Labfors 3 fermentor was used. LB medium containing gentamycin (7 µg/ml) and 0.1% glucose (w/v) was used for fermentation, supplemented with antifoam agent struktol J647 (Schill & Seilacher, Hamburg, Germany). Overnight bacterial culture in LB medium was used for inoculation. In the beginning of fermentation, the pO_2 of dissolved oxygen was 0%, stirring speed 500 rpm and OD_{600} 0.288. Fermentation was performed at 28°C. The level of oxygen was slowly raised to obtain 20% dissolved oxygen (pO_2). A feedback loop to maintain the oxygen level at 20% was set by adjusting the stirring speed ranging from 150 to 1100 rpm. Pumping of the feed solution (50% glycerol, 2.5 g/l $MgSO_4$, 33 ml/h) was initiated at OD~1. The induction of protein production was performed at OD~1.5 by adding 0.25 mM IPTG (Fermentas) and 0.2% (w/v) L–arabinose (Sigma). Induction was continued for 24 h. The cells were harvested by centrifugation (5000×g, 15 min, 20°C), lysed by sonication and subjected to 2-iminobiotin affinity chromatography as previously described [19].

Oligometric state of the produced chimeric dual chain fusions

Size-exclusion chromatography was used to determine the oligomeric state of the produced fusion proteins with an Äkta Purifier P10 instrument (Amersham Biosciences) equipped with a Tricorn™ High Performance Column Superdex™ 75 10/300 GL. A 50-µg sample of protein diluted in liquid phase (50 mM $NaPO_4$ pH 7, 650 mM NaCl) was used. Bio-Rad standard proteins ranging from 1.5 to 67 kDa were used to calibrate the column.

Thermostability of dcAVD/AVR4 protein

For SDS-PAGE thermostability analysis, the purified proteins were first *in situ* acylated [30]. Sulfosuccinimidyl acetate (Pierce #26777) was used for protein acylation. D-biotin (140 µM) was added to half the samples, and an equal volume of water was added to the other half. Samples were diluted with SDS- and 2-mercaptoethanol-containing sample buffer and incubated at different temperatures (without biotin: RT (20–22°C), 40°C, 50°C, 60°C, 70°C, 100°C; with biotin: RT, 70°C, 80°C, 90°C, 100°C) for 20 min. After incubation, the samples were analyzed by SDS-PAGE stained with Coomassie Brilliant Blue. Thermostability of AVR4(C122S) protein was also determined. The temperatures used for AVR4(C122S) protein were without biotin: RT, 60°C, 70°C, 80°C, 90°C, 100°C; and with biotin: RT, 80°C, 90°C, 100°C, 120°C.

The thermodynamics of dcAVD/AVR4 unfolding was characterized by differential scanning calorimetry (DSC) [31]. In the DSC measurements, wt dcAVD was used as a control. The

proteins were analyzed in 50 mM sodium phosphate buffer (pH 7.0) containing 100 mM of sodium chloride. The temperature was increased from 20 to 130°C with a 1°C/min heating rate.

Determination of the dissociation rate constant of fluorescent biotin

The dissociation rate constant of conjugated biotin was examined by fluorescence spectroscopy using fluorescently labeled biotin (Bf560–BTN, Arcdia, Turku, Finland). A 50-nM biotin-fluorochrome solution diluted in 50 mM sodium phosphate buffer (pH 7.0) containing 650 mM of sodium chloride was mixed with the protein solution (50 nM or 100 nM), followed by incubation (5 min). The dissociation process was illuminated by adding 100-fold excess of D-biotin (5 μM), which was measured after 1 hour. The measurements were performed at 25°C or 50°C with QuantaMaster Spectrofluorometer System equipped with thermostated water bath [19].

Determination of 2-iminobiotin binding kinetics by surface plasmon resonance biosensor analysis

The SPR analysis was performed for wt AVD, dcAVD/AVR4 and AVR4 proteins. The CM5 chips functionalized with 2-iminobiotin were used for the interaction analysis. The Biacore X instrument (GE Healthcare) was used and the measurements were performed at 25°C using 50 mM sodium carbonate buffer (pH 11) containing 1 M of sodium chloride as the running buffer. The kinetic data were calculated from the sensograms using Langmuirian binding model implemented in the BIAevaluation® program.

Molecular modeling and molecular dynamics simulations

The predicted models of chimeric dual-chain fusion proteins were generated by exploiting the existing 3-D structures of dcAVD (PDB 2C4I), SA (PDB 1MK5), AVR2 (PDB 1WBI) and AVR4 (PDB 1Y53) as templates. The sequence forming the cp65-subunit of dcAVD (PDB 2C4I) was first replaced with a corresponding sequence from another protein that had its sequence reorganized equally (for details, see Figure 1 and [9]). The program Modeller 9v2 [32] was then used to generate a homology model of the pseudodimer. Structural water molecules were included according to their template structures, and the following structures were also used to position the structural water molecules: PDB 1SLF, PDB 1AVE, and PDB 2AVI. The pseudotetrameric forms were then generated by positioning two pseudodimers by structural superimposition with BODIL [33]. The generated homology models were subjected to molecular dynamics simulation using the program NAMD 2.6 and the CHARMM22 force field [34]. First, the models (including structural waters extracted from PDB-files) were placed in a box filled with TIP3 waters (box size approximately 80 Å×80 Å×70 Å) using the SOLVATE algorithm in the program VMD 1.8.6 [35]. Na^+ or Cl^- ions were added to neutralize the system. The complete systems contained 58377 (dcAVD/AVR4), 56803 (dcAVD/AVR2), and 53207 (dcAVD/SA) atoms. The systems were then subjected to minimization as follows: the first minimization was performed by fixing all the protein atoms and allowing water molecules to move according to minimization procedure implemented in NAMD for 4000 steps. The second 4000 step minimization was performed by releasing all the atoms in the system except Cα atoms. Finally, the system was minimized without constraints for 4000 steps.

The minimized system was thermalized by increasing the temperature to 310 K in 31 ps with a Berendsen barostat (1 atm). This step was followed by the production of the coordinate trajectories under NPT conditions using the Berendsen barostat (1 atm) and the Berendsen thermostat at 310 K. The resulting data were analyzed using the programs VMD 1.8.7, PyMOL 1.3, and MS Excel 2010. RMSF calculations were run such that the Cα atoms of one peptide chain were first aligned, and RMSF (deviation over the last 10 ps) values were calculated for the Cα atoms of the aligned chain. RMSF values were taken every 5 ps for the last 3 ns of equilibration (600 time points) and averaged. The force-field interaction energies between domains of dcAVDs were calculated from the simulation trajectories by the NAMD 2.7 program using 5-ps time step.

Supporting Information

Figure S1 Oligomeric state of the dcAVD/AVR4 protein determined by gel filtration. Gel-filtration chromatography analysis showed a main peak corresponding to a dimeric (pseudo-tetrameric) dcAVD/AVR4 (estimated molecular weight of 46 kDa). Additionally, some higher molecular weight species are detected.
(TIF)

Figure S2 The SDS-PAGE analysis of the pilot-scale production and the 2-iminobiotin purification of the dcAVD/AVR4 protein. A Labfors Infors 3 bioreactor was used for a pilot-scale production of the dcAVD/AVR4 protein. The pilot-scale fed-batch fermentation in standard LB medium yielded over 5 mg of the pure dcAVD/AVR4 protein per liter of production medium. A, chicken avidin (10 μg); T, total sample from culture; L, unbound fraction after incubation with 2-iminobiotin; f2–f17 samples of elution fractions. The PageRuler™ Plus prestained protein ladder (Fermentas) was used as a molecular weight standard.
(TIF)

Table S1 Conditions used in the PCR amplification analysis. In the first PCR experiment, four conditions were used by varying annealing and elongation times. In the second PCR experiment, two different conditions were used as described in the table.
(DOC)

Acknowledgments

The authors acknowledge the contribution of Dr. Henri Nordlund in the early stage of this research. He passed away 21st of July in 2008. The valuable discussions with Dr. Einari Niskanen are acknowledged. We acknowledge the skilled technical support provided by Irene Helkala and Ulla Kiiskinen. CSC – IT Center for Science Ltd is acknowledged for the use of their supercomputer resources for the MD simulations.

Author Contributions

Conceived and designed the experiments: VPH TAR MSK. Performed the experiments: SV TAR SK TKMN JH VPH. Analyzed the data: TAR SK VPH. Contributed reagents/materials/analysis tools: MSK VPH. Wrote the paper: TAR SK SV TKMN MSK VPH.

References

1. Groman EV, Rothenberg JM, Bayer EA, Wilchek M (1990) Enzymatic and radioactive assays for biotin, avidin, and streptavidin. Method Enzymol 184: 208–217.
2. Bayer EA, Ben-Hur H, Wilchek M (1990) Analysis of proteins and glycoproteins on blots. Method Enzymol 184: 415–427.
3. Barbarakis MS, Qaisi WG, Daunert S, Bachas LG (1993) Observation of "hook effects" in the inhibition and dose-response curves of biotin assays based on the interaction of biotinylated glucose oxidase with (strept)avidin. Anal Chem 65: 457–460.

4. Green NM (1970) Spectrophotometric determination of avidin and streptavidin. Method Enzymol 18: 418–424.
5. Lin HJ, Kirsch JF (1979) A rapid, sensitive fluorometric assay for avidin and biotin. Method Enzymol 62: 287–289.
6. Mock DM, Horowitz P (1990) Fluorometric assay for avidin-biotin interaction. Method Enzymol 184: 234–240.
7. Zacco E, Pividori MI, Alegret S (2006) Electrochemical biosensing based on universal affinity biocomposite platforms. Biosens Bioelectron 21: 1291–1301.
8. Lesch HP, Kaikkonen MU, Pikkarainen JT, Ylä-Herttuala S (2010) Avidin-biotin technology in targeted therapy. Expert Opin Drug Deliv 7: 551–564.
9. Nordlund HR, Laitinen OH, Hytönen VP, Uotila ST, Porkka E, et al. (2004) Construction of a dual chain pseudotetrameric chicken avidin by combining two circularly permuted avidins. J Biol Chem 279: 36715–36719.
10. Hytönen VP, Hörhä J, Airenne TT, Niskanen EA, Helttunen KJ, et al. (2006) Controlling quaternary structure assembly: Subunit interface engineering and crystal structure of dual chain avidin. J Mol Biol 359: 1352–1363.
11. Hytönen VP, Nordlund HR, Hörhä J, Nyholm TK, Hyre DE, et al. (2005) Dual-affinity avidin molecules. Proteins 61: 597–607.
12. Leppiniemi J, Määttä JA, Hammaren H, Soikkeli M, Laitaoja M, et al. (2011) Bifunctional avidin with covalently modifiable ligand binding site. PLoS One 6: e16576.
13. Aslan FM, Yu Y, Mohr SC, Cantor CR (2005) Engineered single-chain dimeric streptavidins with an unexpected strong preference for biotin-4-fluorescein. Proc Natl Acad Sci U S A 102: 8507–8512.
14. Keinänen RA, Wallén MJ, Kristo PA, Laukkanen MO, Toimela TA, et al. (1994) Molecular cloning and nucleotide sequence of chicken avidin-related genes 1–5. Eur J Biochem 220: 615–621.
15. Ahlroth MK, Kola EH, Ewald D, Masabanda J, Sazanov A, et al. (2000) Characterization and chromosomal localization of the chicken avidin gene family. Anim Genet 31: 367–375.
16. Weber PC, Cox MJ, Salemme FR, Ohlendorf DH (1987) Crystallographic data for streptomyces avidinii streptavidin. J Biol Chem 262: 12728–12729.
17. Eisenberg-Domovich Y, Hytönen VP, Wilchek M, Bayer EA, Kulomaa MS, et al. (2005) High-resolution crystal structure of an avidin-related protein: Insight into high-affinity biotin binding and protein stability. Acta Crystallogr D Biol Crystallogr 61: 528–538.
18. Hytönen VP, Määttä JA, Kidron H, Halling KK, Hörhä J, et al. (2005) Avidin related protein 2 shows unique structural and functional features among the avidin protein family. BMC Biotechnol 5: 28.
19. Hytönen VP, Laitinen OH, Airenne TT, Kidron H, Meltola NJ, et al. (2004) Efficient production of active chicken avidin using a bacterial signal peptide in *Escherichia coli*. Biochem J 384: 385–390.
20. Laitinen OH, Hytönen VP, Ahlroth MK, Pentikäinen OT, Gallagher C, et al. (2002) Chicken avidin-related proteins show altered biotin-binding and physicochemical properties as compared with avidin. Biochem J 363: 609–617.
21. Gonzalez M, Bagatolli LA, Echabe I, Arrondo JLR, Argarana CE, et al. (1997) Interaction of biotin with streptavidin. thermostability and conformational changes upon binding. The Journal of Biological Chemistry 25: 11288–11294.
22. Sano T, Cantor CR (1990) Cooperative biotin binding by streptavidin. Electrophoretic behavior and subunit association of streptavidin in the presence of 6 M urea. J Biol Chem 265: 3369–3373.
23. Jones ML, Kurzban GP (1995) Noncooperativity of biotin binding to tetrameric streptavidin. Biochemistry 34: 11750–11756.
24. Green NM, Toms EJ (1972) The dissociation of avidin-biotin complexes by guanidinium chloride. Biochem J 130: 707–711.
25. Hytönen VP, Nyholm TK, Pentikäinen OT, Vaarno J, Porkka EJ, et al. (2004) Chicken avidin-related protein 4/5 shows superior thermal stability when compared with avidin while retaining high affinity to biotin. J Biol Chem 279: 9337–9343.
26. Hyre DE, Le Trong I, Merritt EA, Eccleston JF, Green NM, et al. (2006) Cooperative hydrogen bond interactions in the streptavidin-biotin system. Protein Sci 15: 459–467.
27. DeChancie J, Houk KN (2007) The origins of femtomolar protein-ligand binding: Hydrogen-bond cooperativity and desolvation energetics in the biotin-(strept)avidin binding site. J Am Chem Soc 129: 5419–5429.
28. Nordlund HR, Hytönen VP, Hörhä J, Määttä JA, White DJ, et al. (2005) Tetravalent single chain avidin: From subunits to protein domains via circularly permuted avidins. Biochem J 392: 485–491.
29. Airenne KJ, Oker-Blom C, Marjomäki VS, Bayer EA, Wilchek M, et al. (1997) Production of biologically active recombinant avidin in baculovirus-infected insect cells. Prot Exp Pur 9326: 100–108.
30. Bayer EA, Ehrlich-Rogozinski S, Wilchek M (1996) Sodium dodecyl sulfate-polyacrylamide gel electrophoretic method for assessing the quaternary state and comparative thermostability of avidin and streptavidin. Electrophoresis 17370: 1319–1324.
31. Hytönen VP, Laitinen OH, Grapputo A, Kettunen A, Savolainen J, et al. (2003) Characterization of poultry egg-white avidins and their potential as a tool in pretargeting cancer treatment. Biochem J 372: 219–225.
32. Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, et al. (2006) Comparative protein structure modeling using modeller. Curr Protoc Bioinformatics, DOI: 10.1002/0471250953.bi0506s15.
33. Lehtonen JV, Still DJ, Rantanen VV, Ekholm J, Björklund D, et al. (2004) BODIL: A molecular modeling environment for structure-function analysis and drug design. J Comput Aided Mol Des 18: 401–419.
34. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, et al. (2005) Scalable molecular dynamics with NAMD. J Comput Chem 26: 1781–1802.
35. Humphrey W, Dalke A, Schulten K (1996) VMD: Visual molecular dynamics. J Mol Graph 14: 33–8, 27–8.

ACS APPLIED MATERIALS & INTERFACES

RESEARCH ARTICLE

www.acsami.org

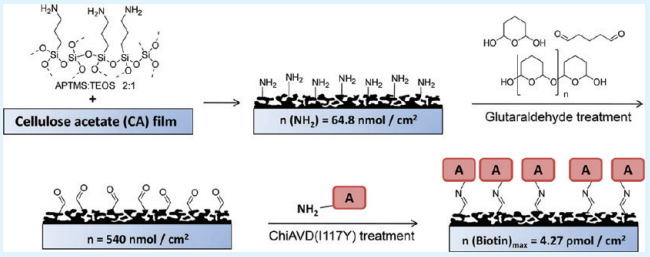
Covalent Biofunctionalization of Cellulose Acetate with Thermostable Chimeric Avidin

Jarkko J. Heikkinen,[†] Tiina A. Riihimäki,[‡] Juha A.E. Määttä,[‡] Sini E. Suomela,[†] Jukka Kantomaa,[†] Markku S. Kulomaa,[‡] Vesa P. Hytönen,[*,‡] and Osmo E.O. Hormi[*,†]

[†]Department of Chemistry, University of Oulu, P.O. BOX 3000, FI-90014 Oulu, Finland

[‡]Institute of Biomedical Technology, University of Tampere and Tampere University Hospital, FI-33014 Tampere, Finland

ABSTRACT: A stable, bioactive cellulose acetate (CA) surface was developed by functionalizing the surface with highly thermostable avidin form. The CA films were first functionalized with a mixture of 3-aminopropyltrimethoxysilane and tetraethoxysilane to introduce free amino groups onto the surface of CA films. Free amino groups were functionalized with glutaraldehyde to obtain an activated surface for covalent biomolecule immobilization. A genetically engineered, high-affinity biotin-binding protein chimeric avidin, ChiAVD(I117Y), was used for biofunctionalization of the surface. The chimeric avidin protein has an increased stability in chemically harsh conditions and at high temperature when compared to wt (strept)avidin. The biological activity, i.e., biotin-binding capacity, of the immobilized protein was probed by [^{3}H]-biotin. The activity of the chimeric avidin on functionalized CA films was fully retained over the three months' study period. The biotin-binding capacity of the immobilized chimeric avidin was compared to that of immobilized streptavidin, chicken avidin, and rhizavidin and significant differences between proteins were detected. The developed material offers a valuable platform for the development of inexpensive in vitro diagnostics and also supports biosensing applications that are required to operate under demanding conditions.



KEYWORDS: cellulose diacetate, sol−gel functionalization, glutaraldehyde, chimeric avidin, protein immobilization, thermostable protein, biomaterials, biotin

1. INTRODUCTION

The interest in protein immobilization techniques for biochips and biosensors has grown rapidly. A number of reviews have been published on the immobilization of proteins.[1−7] However, despite extensive studies, it is still challenging to immobilize proteins onto surfaces in such a way that their three-dimensional structure, functionality, and binding sites are well retained. One way to overcome this challenge is to use a tetrameric avidin containing four identical binding sites for biotin as an adaptor protein. The avidin-functionalized surface is used in many life science applications as a bioaffinity capturing tool for controlled immobilization of biotinylated protein.[8] This is based on avidins ultratight affinity toward the H-vitamin, also called as biotin ($K_d \approx 1 \times 10^{-15}$ M).[9] Avidin belongs to the chicken avidin family that also contains several related proteins including avidin related protein 4 (AVR4). AVR4 is one of the most thermally stable proteins among the protein family and in our previous study the stabilizing structural elements of AVR4 were transferred to wt avidin. The resulting chimeric avidin mutant ChiAVD(I117Y) (hereinafter referred to as chimeric avidin) was found to possess significantly increased thermal and chemical stabilities as compared to those of wt (strept)avidin.[10,11] Another valuable property of this protein is that it can be inexpensively produced in *E. coli* in a functionally active form. Because of these favorable characteristics the use of chimeric avidin for the immobilization may enable cost-effective manufacturing of biochips and biosensors that remain stable during long periods of storage or under demanding production conditions.

The main objective of this research was the development of a method for covalent immobilization of chimeric avidin on a cellulose acetate (CA) surface. CA, also known as zyl or zylonite, is valuable material utilized widely in manufacturing of films and fibers. A sol−gel-based method was chosen to first introduce free amino groups onto CA film using mixture of 3-aminopropyltrimethoxysilane (APTMS) and tetraethoxysilane (TEOS). This amino-functionalized CA film was then reacted with glutaraldehyde to form an activated CA film, which could then be used directly for covalent protein immobilization. The biotin-binding capacity of immobilized chimeric avidin was determined using tritium-labeled biotin ([^{3}H]-biotin). We also studied the durability of the surfaces functionalized with different methods over a three-month period. Finally, a comparative study was performed using other biotin-binding proteins chicken avidin, streptavidin and rhizavidin as linker proteins.[11]

2. EXPERIMENTAL SECTION

2.1. Materials. Cellulose diacetate film (CA) was obtained from Clarifoil (P27, thickness 150 μm). 3-Aminopropyltrimethoxysilane (APTMS), tetraethoxysilane (TEOS), glutaraldehyde (25% (w/v) in

Received: October 29, 2010
Accepted: May 25, 2011

   dx.doi.org/10.1021/am200272u | ACS Appl. Mater. Interfaces XXXX, XXX, 000–000

H_2O) and bovine serum albumin (BSA) were purchased from Aldrich and used without further purification.

Chimeric avidin and rhizavidin core were produced in BL21-AI *E. coli* cells by using pilot-scale fermentor and purified with 2-iminobiotin affinity chromatography as described previously.[12] Chicken avidin was purchased from Belovo (Bastogne, Belgium) and streptavidin from Calbiochem (189730−10MG). D-biotin was purchased from Sigma (B-4501) and D-[8,9-^{3}H]-biotin from Amersham (Buckinghamshire, England).

2.2. Functionalization and Characterization of CA Films. *Amino-Functionalization of CA Film.* To prepare a sol−gel stock solution for CA film amino-functionalization, we dissolved TEOS (0.75 mL) in ethanol (4 mL) and 10 μL of 4 M hydrochloric acid was added. In a separate flask, APTMS (1.5 mL) was dissolved in ethanol (4 mL) and 4 mL of water was added. Both solutions were stirred for 30 min after TEOS solution was added to the APTMS solution. This mixture was stirred for 15 min and then diluted 5-fold by adding 48 mL of ethanol. The obtained stock solution was stored at fridge (+4 °C). After the dilution of the stock solution by ethanol (1:5) CA films were dip-coated by immersing the film to the sol−gel solution for 2 min. The functionalized CA films were dried at room temperature (RT, 21−23 °C) and stored in an desiccator prior the use.

Glutaraldehyde Functionalization. Amino-functionalized CA slides were placed in a glutaraldehyde solution (10% in H_2O) for 1 h at RT. After glutaraldehyde treatment, the CA slides were washed with water, dried, and stored in an desiccator at RT prior the use.

CA Film Characterization. Gas chromatography was used to measure the density of free amino groups on the amino functionalized CA film and amino-reactive groups on the glutaraldehyde functionalized CA film.[13] The amino and amino-reactive groups on the surface of functionalized films were reacted with benzaldehyde and aniline analytes, respectively, to form imine bonds when refluxed in anhydrous toluene solution for 4 h. Anisole was used as a standard in the GC measurements to calculate the amount of reacted analyte with the functionalized surface. The amounts of amino and amino-reactive groups were calculated per area of 1 cm^2. Scanning electron microscope (SEM) images of prepared CA films were analyzed using field-emission scanning electron microscope Zeiss ULTRA plus.

2.3. Immobilization of Avidins onto the CA Films. Films were cut to 1 cm × 1 cm size and placed into the wells of 24-well plate. A 200 μL aliquot of avidin solution (0.1 mg/mL in PBS: 0.137 M sodium chloride, 2.8 mM potassium chloride, 11.9 mM Na-phosphate buffer, pH 7.4) was added to each well and films were incubated in the avidin solution for one hour at 37 °C. The films were washed three times with 300 μL of PBST buffer (PBS with 0.05% of TWEEN 20) and blocked with 3% BSA solution in PBS for one hour. The films were finally washed five times with 300 μL of PBST and dried at RT. The avidin functionalized films were either stored at RT or in a fridge at +4 °C prior the biotin-binding capacity measurements. Following abbreviations are used for the films studied: 1 = unmodified CA film, 2 = amino-functionalized CA film, 3 = glutaraldehyde functionalized CA film, A = chimeric avidin coated film, wtA = wt avidin coated film, RA = rhizavidin coated film, SA = streptavidin coated film, and R = control films without avidin coating.

2.4. Determination of Biotin-Binding Capacity with [^{3}H]-Biotin. The biotin-binding capacity of surface-immobilized avidin layer was measured using D-[8,9-^{3}H]biotin thereafter referred to as [^{3}H]-biotin. Nonspecific binding of the biotin was taken into account by measuring the bound [^{3}H]-biotin after blocking the films with an excess of D-biotin. These results were subtracted from the results without D-biotin blocking to calculate specific biotin-binding capacity. All biotin-binding capacity results are calculated as a bound biotin per area of the film [mol/cm^2].

General [^{3}H]-Biotin Assay Protocol. The CA films (1 × 1 cm) were incubated for 2 h at RT in a 1 mL sodium phosphate solution (50 mM NaH_2PO_4/Na_2HPO_4, 100 mM NaCl, pH 7.0) to wash the films. The solution was removed and 1 mL of 12.8 nM [^{3}H]-biotin solution (50 mM NaH_2PO_4/Na_2HPO_4, 100 mM NaCl, 10 μg/mL BSA, pH 7.0) was added and incubated for 1 h at RT. A sample (90 μL) was taken from the solution and added to 4 mL of Opti-Phase scintillation solution (PerkinElmer). The scintillations were counted with LKB Wallac 1217 Rackbeta liquid scintillation counter (Turku, Finland). The amount of bound [^{3}H]-biotin was calculated based on counts measured from free [^{3}H]-biotin solution before used for the experiment.

[^{3}H]-Biotin Assay Protocol with D-Biotin Blocking. The CA films were washed first with 1 mL of sodium phosphate -solution for 1 h. The solution was removed and 1 mL of D-biotin solution (50 mM NaH_2PO_4/Na_2HPO_4, 100 mM NaCl, pH 7.0 with 150 μM D-biotin) was added and the films were incubated in this solution for one hour. The D-biotin solution was removed and [^{3}H]-biotin assay was then carried out as described above.

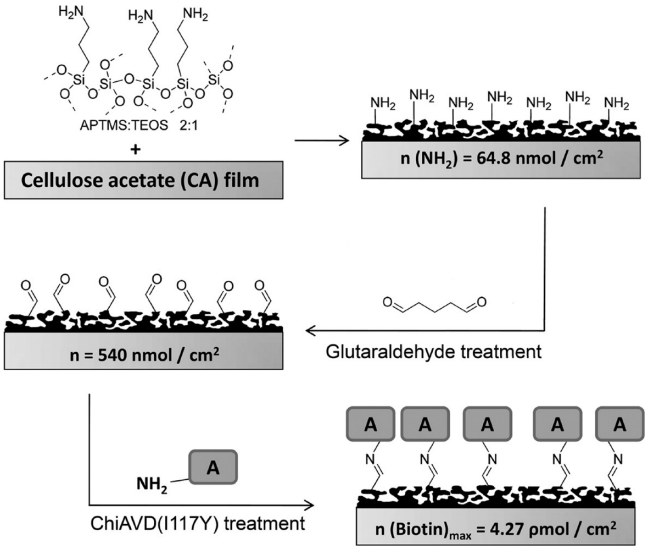


Figure 1. Schematic representation of functionalization and protein immobilization on the CA surface. For clarity, the chemistry of glutaraldehyde on the surfaces of CA films is shown as a monomeric dialdehyde.

3. RESULTS AND DISCUSSION

3.1. Functionalization of CA Films. The functionalization was done in two steps (Figure 1). During the first step, the CA films were treated with a sol−gel solution made from APTMS and TEOS in ethanol to functionalize CA films with free amino groups. A ratio of 2:1 between APTMS and TEOS were found to yield stable amino-functionalized CA film surface.[14] In the second step, a glutaraldehyde treatment was done by stirring the amino-functionalized CA films in a 10% glutaraldehyde solution in water for 1 h. During the glutaraldehyde treatment, the CA slides turned slightly red, indicating a reaction between glutaraldehyde and amino groups.

To measure the aging effect of sol−gel stock solution, a series of functionalized films were prepared where the sol−gel stock solution aging time varied from zero days to ten days. Determination of the free amino groups by gas chromatography revealed that seven days old sol−gel stock solution gave the highest amount of amino groups (64.8 nmol/cm^2) after amino-functionalization (Figure 2). However, all films regardless of the aging time clearly had a very high amount of amino groups (26−65 nmol/cm^2). In the SEM analysis, a rough surface structure with high surface area was observed (Figure 3). This can explain the high amino content on the R2 films. Similar high amino densities have been reported for amino-functionalized poly(ethylene-*co*-acrylic acid) polymer films, which were functionalized by using amino-modified silica nanoparticles.[15]

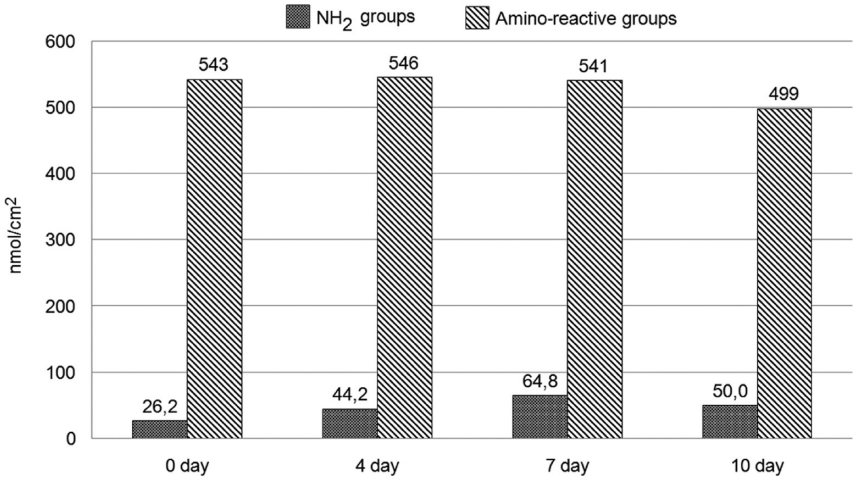


Figure 2. Density of free amino groups on R2 surface [nmol/cm^2] and amino-reactive groups on R3 surface as a function of the sol−gel aging time.

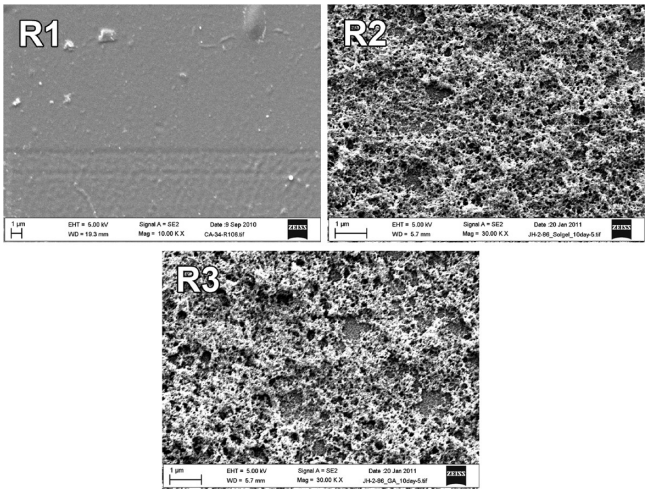


Figure 3. Scanning electron microscopy analysis of functionalized CA films. R1 = unmodified CA film, R2 = sol−gel amino-functionalized CA film, R3 = glutaraldehyde functionalized CA film.

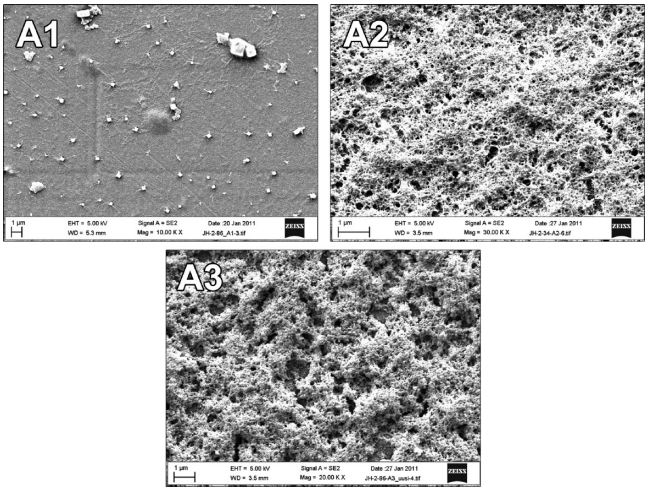


Figure 4. Scanning electron microscopy analysis of CA films after incubation with chimeric avidin. A1 = unmodified CA film treated with chimeric avidin, A2 = amino functionalized CA film treated with chimeric avidin, A3 = glutaraldehyde functionalized CA film treated with chimeric avidin.

All prepared glutaraldehyde-functionalized R3 films were found to have roughly ten times more amino-reactive groups (500−560 nmol/cm^2) compared to free amino groups after amino functionalization. Several studies have shown that commercially available glutaraldehyde−water solution represents multicomponent mixtures including aldehydes, monohydrates, dihydrates, cyclic hemiacetals, and oligomers and polymers of these.[16] All these forms can react with amino groups, which can explain the high density of amino-reactive groups on the surface of R3 films. Also, because the commercial glutaraldehyde can exist as a polymer form, especially when stored at RT, a low density of free amino groups before glutaraldehyde functionalization (26.2 nmol/cm^2, 0 day) resulted in the same level of amino-reactive groups after GA functionalization when compared to other films with higher (44−65 nmol/cm^2) amino group density.

3.2. Immobilization of Avidin Proteins. The immobilization of avidin proteins (chimeric avidin, chicken avidin, rhizavidin, streptavidin) onto the CA films was performed by incubating the CA films in a solution containing 0.1 mg/mL avidin protein for one hour at 37 °C (Figure 1). Thorough washing with PBST buffer was effective in the removal of the weakly bound avidin proteins from the surface. Measured SEM images of chimeric avidin immobilized films after PBST washing period are shown in Figure 4, indicating the stability of amino and glutaraldehyde functionalization after protein immobilization. F4

3.3. Biotin-Binding Capacity. The specific biotin-binding capacity of the films was determined by comparing the measured total biotin-binding capacity with the results obtained after blocking the surfaces with D-biotin before [^{3}H]-biotin assay. No significant differences were observed between different control films (R1-R3) in their binding capacities.

All chimeric avidin-coated films showed high biotin-binding capacity. Initially (the first day) both nonfunctionalized film (A1) and glutaraldehyde activated film (A3) bound almost the same amount of [^{3}H]-biotin, indicating that the same amount of chimeric avidin was immobilized onto both films. The amino-functionalized film (A2) was found to have one-half less immobilized chimeric avidin compared to A1 and A3 films.

dx.doi.org/10.1021/am200272u |ACS Appl. Mater. Interfaces XXXX, XXX, 000–000

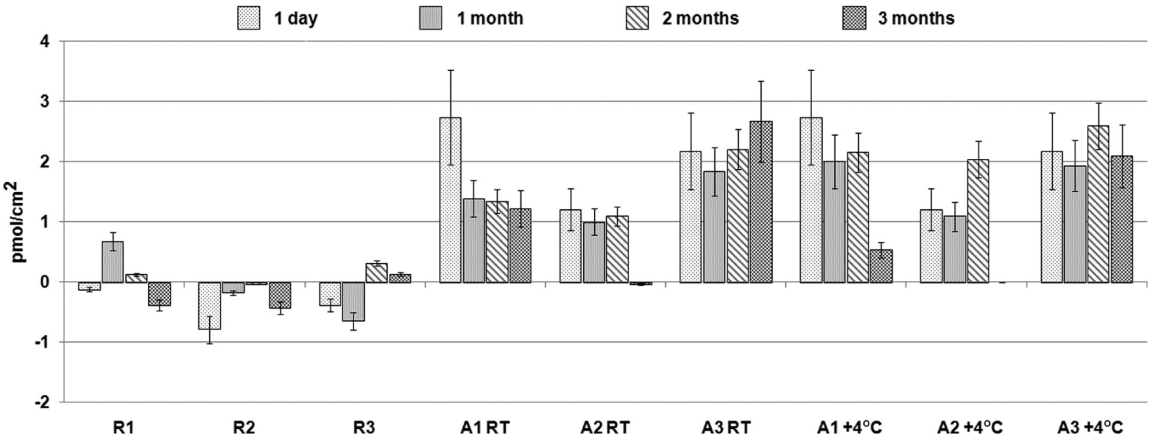


Figure 5. Botin-binding capacity of the CA films. The control films (R1, R2, R3) showed no biotin-binding activity. The most stable film was the sol−gel- and glutaraldehyde-functionalized CA film with chimeric avidin coating (A3), because the binding activity of that film did not decrease during the 3-month storage period. Furthermore, no difference was observed between the different storage conditions in the case of A3. The A1 film stored at RT already showed a significant decrease in binding capacity after one month of storage, whereas the film stored at +4 °C showed slower decay of the binding capacity. (The A2 sample stored at +4 °C for 3 months was not analyzed.).

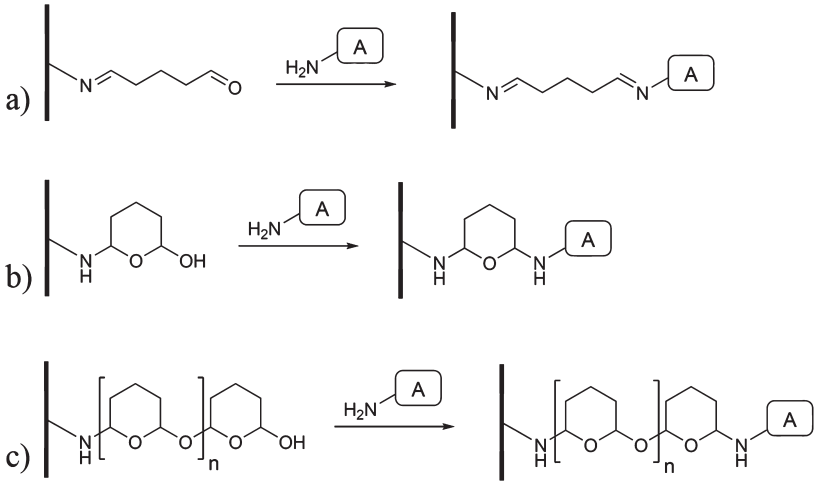


Figure 6. Reactions of different glutaraldehyde forms with protein (A) under neutral conditions: (a) glutaraldehyde, (b) cyclic hemiacetal, (c) polymeric hemiacetal.

During the 3 month study period, the A3 film was found to be the most stable: covalently bound avidin remained fully active. Additionally, there were no significant differences in binding capacities between A3 samples stored at different temperatures. The biotin-binding activity of A1 film stored at RT already showed a significant decrease after one month's storage, whereas the decay was slower when the film was stored at +4 °C (Figure 5). The difference between the behavior of A3 and A1 films can be explained by the means of protein attachment to the films. In the case of A1 film, avidin is physisorbed on the film, whereas in the case of A3 film the protein is covalently bound to the surface. It is well-known that glutaraldehyde has preservative properties against biomolecular unfolding, since the covalent multipoint attachment, presumed to occur with glutaraldehyde, prevents the unfolding of proteins.[17] Moreover, the polymeric nature of glutaraldehyde provides a long leash, attaching the protein to the matrix, which may permit greater flexibility for protein conformational changes required for activity. Under the used neutral conditions used for protein immobilization, a reaction between cyclic hemiacetal or its polymeric form and an amino group in protein is reported to form the most stable bond,[18] but several different reactions can proceed simultaneously leading to protein immobilization (Figure 6). Amino-functionalized film (A2) lost biotin-binding capacity completely after three months when stored at room temperature. Unfortunately, we were not able to analyze the sample of A2 stored at +4 °C after 3 months, so the difference between storage conditions can not be evaluated.

The binding capacity of chimeric avidin coating can also be evaluated based on molecular dimensions. According to the 3D-structure,[12] chimeric avidin can be described as a box with a dimension of about 5 nm, and a full monolayer of the avidins would result in a protein density of 6.6 pmol/cm^2. The measured specific biotin-binding capacity of A3 surface was 2.17 pmol/cm^2 (1 day). Therefore, the generated surface had a capacity that is clearly lower than a fully active avidin monolayer, and, by converting the obtained values for molecular density, there were 0.33 biotin-binding sites per surface area of an avidin tetramer. Based on the crystal structure, (chimeric) avidin is quite symmetric. Therefore, the number of available binding sites on

dx.doi.org/10.1021/am200272u |ACS Appl. Mater. Interfaces XXXX, XXX, 000–000

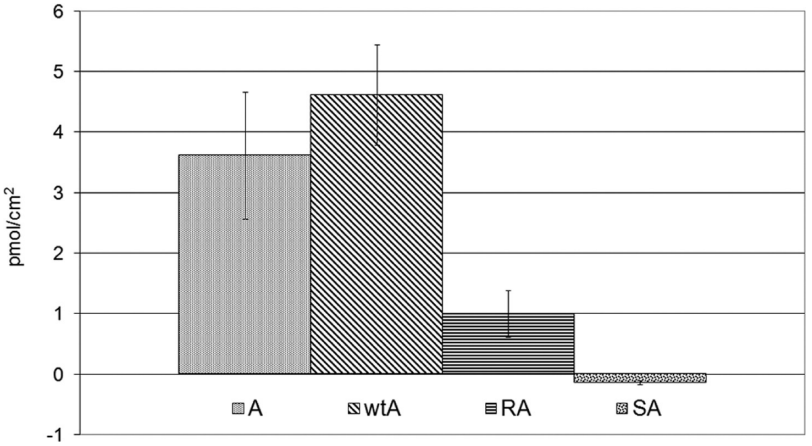


Figure 7. Comparison of biotin-binding capacity of the activated CA surfaces coated with various different avidins. Chimeric avidin (A) was found to behave similarly to chicken avidin (wtA), whereas in the case of rhizavidin (RA) and streptavidin (SA), a low binding capacity was observed.

surface-immobilized avidin would be two or less, and only one binding site per immobilized protein would be a fair assumption. Therefore, in the case of A3 films, the determined specific binding capacity suggests an average coverage of approximately 33% or less of the surface with functional chimeric avidins. In reality, because of the roughness of the surface (Figures 3 and 4), the coverage may be significantly smaller.

We also analyzed films coated with chicken avidin, streptavidin and rhizavidin. Interestingly, chimeric avidin and chicken avidin showed similar high biotin-binding capacities, whereas films coated with streptavidin and rhizavidin had low biotin-binding capacities (Figure 7). This may reflect the chemical reactivity of the proteins with the activated surface. Other possible explanations are the physicochemical differences between proteins. Rhizavidin is dimeric avidin form, which has significantly lower thermal stability when compared to chicken avidin.[11] Streptavidin in turn, has thermal stability similar to chicken avidin, but appears to be more sensitive to treatment with different chemicals, such as methanol.[12] Therefore, it is possible that the result reflects the chemical stress applied to the immobilized protein because of the adjacent surface. Chimeric avidin was selected for the experiment instead of wt avidin, due to its more stable structure in harsh conditions. Also, the efficient and scalable expression in *E. coli* supported the use of chimeric avidin in the study. However, the long-term stability of wt avidin on the CA films should be studied more closely in the future.

4. CONCLUSION

A new method for biofunctionalization of cellulose acetate films by a simple two-step protocol was developed. A sol−gel solution containing 3-aminopropyltrimethoxysilane and tetraethoxysilane was first used to introduce free amino groups onto the surface of cellulose acetate (CA) film. The amino groups were then reacted with glutaraldehyde to obtain a reactive glutaraldehyde layer on the surface of CA film. The protein immobilization was studied with highly thermostable chimeric avidin, which is a genetically engineered version of the high-affinity biotin-binding protein avidin. The activity of CA-immobilized chimeric avidin retained completely for three months when stored at RT or at +4 °C when assayed by a tritium-labeled biotin assay. Immobilization of other avidins (streptavidin, chicken avidin, and rhizavidin) was also analyzed and chimeric avidin and avidin were found to behave similarly. In contrast, streptavidin and rhizavidin yielded notably weaker biotin-binding capacities. The method developed in the study opens up a great potential for the use of avidin-coated CA-films as universal base for various applications, for example, in the development of an inexpensive and sensitive diagnostic tools to be used in personalized medicine platforms.

AUTHOR INFORMATION

Corresponding Author

*E-mail vesa.hytonen@uta.fi (V.P.H.); osmo.hormi@oulu.fi (O. E.O.H.). Phone +358-40-1901517 (V.P.H.); +385-8-5531631 (O.E.O.H.).

ACKNOWLEDGMENT

This project was supported by TEKES−the Finnish Funding Agency for Technology and Innovation (BioFace 40055/08). J.J. H. thanks the Faculty of Science at the University of Oulu and Oulu University Scholarship Foundation for financial support. T. A.R. received grants from Tampere Graduate Program in Biomedicine and Biotechnology (TGPBB). We thank Pirkanmaa Hospital District for financial support. We thank John Saeger for proofreading the manuscript.

REFERENCES

(1) Rusmini, F.; Zhong, Z.; Feijen, J. *Biomacromolecules* **2007**, *8*, 1775–1789.
(2) Wong, L. S.; Khan, F.; Micklefield, J. *Chem. Rev.* **2009**, *109*, 4025–4053.
(3) Jonkheijm, P.; Weinrich, D.; Schroeder, H.; Niemeyer, C. M.; Waldmann, H. *Angew. Chem., Int. Ed.* **2008**, *47*, 9618–9647.
(4) Cretich, M.; Damin, F.; Pirri, G.; Chiari, M. *Biomol. Eng.* **2006**, *23*, 77–88.
(5) Tomizaki, K.; Usui, K.; Mihara, H. *Chembiochem* **2005**, *6*, 782–799.
(6) Lin, P.; Weinrinch, D.; Waldmann, H. *Macromol. Chem. Phys.* **2010**, *211*, 136–144.
(7) Wilson, D. S.; Nock, S. *Curr. Opin. Chem. Biol.* **2002**, *6*, 81–85.
(8) Wilchek, M.; Bayer, E. A. *Biomol. Eng.* **1999**, *16*, 1–4.
(9) Green, N. M. *Adv. Protein Chem.* **1975**, *29*, 85–133.
(10) Hytönen, V. P.; Määttä, J. A. E.; Nyholm, T. K.; Livnah, O.; Eisenberg-Domovich, Y.; Hyre, D.; Nordlund, H. R.; Hörhä, J.;

330 Niskanen, E. A.; Paldanius, T.; Kulomaa, T.; Porkka, E. J.; Stayton, P. S.;
331 Laitinen, O. H.; Kulomaa, M. S. *J. Biol. Chem.* **2005**, *280*, 10228–10233.
332 (11) Helppolainen, S. H.; Nurminen, K. P.; Määttä, J. A.; Halling,
333 K. K.; Slotte, J. P.; Huhtala, T.; Liimatainen, T.; Ylä-Herttuala, S.;
334 Airenne, K. J.; Narvanen, A.; Jänis, J.; Vainiotalo, P.; Valjakka, J.;
335 Kulomaa, M. S.; Nordlund, H. R. *Biochem. J.* **2007**, *405*, 397–405.
336 (12) Määttä, J. A.; Eisenberg-Domovich, Y.; Nordlund, H. R.;
337 Hayouka, R.; Kulomaa, M. S.; Livnah, O.; Hytönen, V. P. *Biotechnol.*
338 *Bioeng.* **2011**, *108*, 481–490.
339 (13) Mehdi, A.; Reyé, C.; Brandés, S.; Guilard, R.; Corriu, R. J. P.
340 *New J. Chem.* **2005**, *29*, 965–968.
341 (14) Yang, G.; Wu, J.; Xu, G.; Yang, L. *Colloids Surf., B.* **2010**,
342 *78*, 351–356.
343 (15) Scaffaro, R.; Botta, L.; Lo, Re, G.; Bertani, R.; Milani, R.; Sassi,
344 A. *J. Mater. Chem.* **2011**, *21*, 3849–3857.
345 (16) Migneault, I.; Dartiguenave, C.; Bertrand, M. J.; Waldron, K. C.
346 *BioTechniques* **2004**, *37*, 790–802.
347 (17) Scouten, W. H.; Luong, J. H. T.; Brown, R. S. *Trends Biotechnol.*
348 **1995**, *13*, 178–185.
349 (18) Walt, D. R.; Agayn, V. I. *TrAC, Trends Anal. Chem.* **1994**,
350 *13*, 425–430.

BMC Biotechnology

BioMed Central
The Open Access Publisher

Modification of the loops in the ligand-binding site turns avidin into a steroid-binding protein

BMC Biotechnology 2011, **11**:64 doi:10.1186/1472-6750-11-64

Tiina A Riihimaki (tiina.riihimaki@uta.fi)
Soili Hiltunen (soili.hiltunen@uta.fi)
Martina Rangl (Martina.Rangl@jku.at)
Henri R Nordlund (henri.nordlund@uta.fi)
Juha AE Maatta (juha.maatta@uta.fi)
Andreas Ebner (Andreas.Ebner@jku.at)
Peter Hinterdorfer (peter.hinterdorfer@jku.at)
Markku S Kulomaa (markku.kulomaa@uta.fi)
Kristiina Takkinen (kristiina.takkinen@vtt.fi)
Vesa P Hytonen (vesa.hytonen@uta.fi)

ISSN 1472-6750

Article type Research article

Submission date 29 April 2011

Acceptance date 9 June 2011

Publication date 9 June 2011

Article URL http://www.biomedcentral.com/1472-6750/11/64

Modification of the loops in the ligand-binding site turns avidin into a steroid-binding protein

Tiina A. Riihimäki[1], Soili Hiltunen[1], Martina Rangl[2], Henri R. Nordlund[1†], Juha A.E. Määttä[1], Andreas Ebner[2], Peter Hinterdorfer[2], Markku S. Kulomaa[1], Kristiina Takkinen[3] and Vesa P. Hytönen[1§]

[1] Institute of Biomedical Technology, University of Tampere and Tampere University Hospital, FI-33520 Tampere, Finland

[2] Institute of Biophysics, Johannes Kepler University Linz, 4040 Linz, Austria

[3] VTT Technical Research Centre of Finland, FI-02044 VTT, Finland

[†]Deceased on July 21st, 2008

[§]To whom correspondence should be addressed: tel +358 401901517 fax +358 335517710 email vesa.hytonen@uta.fi

Authors' email addresses:

TAR: tiina.riihimaki@uta.fi
SH: soili.hiltunen@uta.fi
MR: martina.rangl@jku.at
HRN: deceased, no email address
JAEM: juha.maatta@uta.fi
AE: andreas.ebner@jku.at
PH: peter.hinterdorfer@jku.at
MSK: markku.kulomaa@uta.fi
KT: kristiina.takkinen@vtt.fi
VPH: vesa.hytonen@uta.fi

Abstract

Background: Engineered proteins, with non-immunoglobulin scaffolds, have become an important alternative to antibodies in many biotechnical and therapeutic applications. When compared to antibodies, tailored proteins may provide advantageous properties such as a smaller size or a more stable structure.

Results: Avidin is a widely used protein in biomedicine and biotechnology. To tailor the binding properties of avidin, we have designed a sequence-randomized avidin library with mutagenesis focused at the loop area of the binding site. Selection from the generated library led to the isolation of a steroid-binding avidin mutant (sbAvd-1) showing micromolar affinity towards testosterone (K_d ~ 9 μM). Furthermore, a gene library based on the sbAvd-1 gene was created by randomizing the loop area between *β*-strands 3 and 4. Phage display selection from this library led to the isolation of a steroid-binding protein with significantly decreased biotin binding affinity compared to sbAvd-1. Importantly, differential scanning calorimetry and analytical gel-filtration revealed that the high stability and the tetrameric structure were preserved in these engineered avidins.

Conclusions: The high stability and structural properties of avidin make it an attractive molecule for the engineering of novel receptors. This methodology may allow the use of avidin as a universal scaffold in the development of novel receptors for small molecules.

Keywords: protein engineering, avidin scaffold, phage display, steroid hormone, testosterone

Background

Antibodies are the most widely used biomolecules for therapeutic, diagnostic and research applications, because they can be generated against virtually any molecule using protein engineering techniques (for a review see [1]). However, antibodies have certain fundamental disadvantages such as the complex architecture of their antigen-binding site, low stability, and a rather large size [2-5]. Moreover, the production of full-size antibodies is relatively expensive [6]. To overcome these limitations and to improve therapeutic antibodies, antibodies have been extensively engineered [7]. For example the size of the antibody molecule has been reduced by producing single-domain antigen-binding derivatives [8]. In addition to extensive antibody engineering, a versatile repertoire of tailored biomolecules from non-immunoglobulin protein

scaffolds have been generated [4, 9, 10]. Anticalins, derived from the lipocalin fold, are a good example of engineered proteins [11]. The β-barrel structure of lipocalins is thermostable and robust and serves as an excellent scaffold for engineering novel receptors. They have been modified to bind novel ligands, such as fluorescein and digoxigenin, with affinities comparable with antibodies [12].

Chicken avidin (Avd), known for its extremely high affinity towards the water-soluble vitamin H, D-biotin, has been widely used in life science research applications [13]. Aside from biotin, Avd also binds dyes and peptides, which share no significant structural similarity with biotin [14, 15]. Avd provides an attractive robust scaffold for the development of novel receptors, and Avd has many advantageous properties such as high chemical and thermal stability, a deep ligand binding site optimized for the binding of small molecules, and an oligomeric nature enabling signal amplification. Moreover, the structure of Avd is well characterized [16, 17] and numerous engineered forms of Avd have been described [18]. Engineered Avd forms, in which the two pairs of the binding sites (dual-chain Avd) [19, 20, 21] or all four binding sites (single-chain Avd) [22] can be independently manipulated, have been developed.

In the present study Avd was modified to bind steroid hormones. Avd mutant sbAvd-1, which was captured using the phage display selection [23], has a micromolar affinity to testosterone. The steroid-binding protein sbAvd-1 was characterized and further engineered to decrease cross-reactivity towards other molecules, especially towards Avd's natural ligand biotin. The resulting sbAvd-2 mutant with a modified loop between β-strands 3 and 4 was found to prefer steroid hormones over biotin in ligand binding.

Results

Functional display of Avd protein on the M13 phage

Avd was displayed on the surface of the M13 phage as a fusion with the C-terminal region of the minor coat protein pIII. Two different strategies for displaying the Avd scaffold in the active form on the M13 phage were evaluated. In the first display construct, Avd was produced solely as a fusion with pIII (Avd-pIII; Figure 1A), whereas in the Avd/Avd-pIII display construct, free Avd subunits were produced in addition to the pIII fusion (Figure 1B).

The size and oligomeric state of the Avd-pIII fusions were analyzed by SDS-PAGE and by western blot. Polyclonal anti-Avd (Figure 2A) indicated expression of both the Avd-pIII fusion (upper arrow, ~38 kDa) and the free Avd (lower arrow, ~15 kDa). The production of free subunits should enhance the functional assembly of the tetrameric Avd scaffold, especially if membrane anchoring of Avd by the pIII fusion partner has a negative effect on the oligomerization of Avd subunits. This strategy mimics the generally used amber-stop codon technique, in which free subunits are produced by the read-through of amber stop codon by the tRNA. We assume that the oligomerization of Avd, secreted by the *pelB* signal sequence, occurs in the periplasmic space of *E. coli,* allowing the display of functional Avd on the phage, as is the case for the display of the antibody Fab fragment, which requires folding of the heavy and light chains for the assembly of a functional antibody molecule [24].

A portion of the Avd-pIII fusion was partially proteolytically cleaved, as can be seen from the blot analyzed with the anti-pIII monoclonal antibody (Figure 2B, ~35 kDa band). Because Avd display constructs designed in this study were based on the monovalent display mode (3+3), the intact pIII (migrating at ~58 kDa) expressed from the helper phage was also detected from the blot (Figure 2B, uppermost arrow).

Avd-displaying phages were functional because they bound specifically to the biotin-coated surfaces, and phages were efficiently amplified even after several panning rounds. Moreover, phagemid DNA with the Avd display expression unit, which was confirmed by the restriction enzyme digestion analysis, was stable during the panning rounds (data not shown).

To analyze the functionality of phages, the mixture of amplified Avd and Avd mutant N118M phages were screened by panning against 4-hydroxyazobenzene-2-carboxylic acid (HABA). As determined in our previous study, the biotin binding affinity of the Avd mutant N118M was reduced ~1,000,000-fold ($K_d = 4.2 \times 10^{-9}$ M) and HABA-affinity was increased ~1.5-fold ($K_d = 5.2 \times 10^{-6}$ M) compared to wtAvd [25] (Additional file 1). During the panning procedure a clear enrichment of Avd(N118M) phages over wt Avd phages was detected, which was an indication of the high selectivity of the produced phages. After only three rounds of selection phages displaying the Avd mutant N118M outcompeted the wtAvd phage population (Additional file 2).

Capture of a steroid-binding avidin

The loops adjacent to the ligand-binding site were selected for randomization [16]. First we created a library (Avd L1,2 library) in which residues N12, D13, L14, G15, and S16 were randomized. These amino acids form a loop between β-strands 1 and 2 (Figure 3). Codon NNN was used for randomization, and therefore, all 20 different amino acids were present, including all stop codons. The Avd L1,2 library was ligated into a phagemid as a sole fusion with the C-terminal portion of pIII. Based on sequencing results and transformation efficiency, the Avd L1,2 library was found to consist of approximately 1×10^5 individual members. When a stretch of five amino acid residues is completely randomized, the theoretical library size is 3.2×10^6.

In the panning experiments Avd L1,2 loop library phages were introduced onto a testosterone-coated surface. The phage genomes carrying the mutated cDNA were found to be stable during the selection. A clear enrichment of sequences in the randomized loop area was observed, indicating the success of the selection conditions. Interestingly, we detected N12 as being a highly conserved amino acid residue among the enriched pool of proteins. An Avd variant, named sbAvd-1, that carried the sequence N12, R13, M14, N15, H16 was selected for further analysis.

The specificity of sbAvd-1 can be tuned by additional mutations in the loop between β-strands 3 and 4

To further lower the biotin-binding affinity and to decrease the cross-reactivity of steroid-binding Avd, a library (sbAvd-1 L3,4) was generated in which the loop area between *β*-strands 3 and 4 was randomized (Figure 3). The loop between *β*-strands 3 and 4 is highly important for biotin binding of Avd because this loop 'locks' biotin into the binding site. In this process, three amino acid residues in the loop form direct interactions with biotin [16]. In the library four amino acids (T35, A36, V37 and, T38) were randomized using the NNY codon. The use of this codon covers 14 of the 20 naturally occurring amino acids while eliminating all of the stop codons. The library consisted of approximately 1.4×10^6 individual members, when calculated from the sequencing results and transformation efficiency, which exceeds the theoretical size of the library calculated based on the possible combinations of amino acid residues (3.8×10^4).

Binders from the sbAvd-1 L3,4 library were selected by phage display panning against a testosterone surface. In every panning round washes were adjusted according to the number of output colonies. The quality of the phage genomes carrying the mutated cDNA was evaluated by DNA sequencing at various stages during selection. A combination of acid and testosterone was used for elution. The selected sbAvd-1 phage clones were evaluated by microplate analysis using

BSA-testosterone as a target ligand and utilizing M13-antibody to determine the amount of bound phages. The sbAvd-1 variant that showed the highest binding activity in the microplate assay (data not shown) had the sequence A35, T36, V37, N38. This mutant, named sbAvd-2 was selected for comparative analysis with sbAvd-1.

Production and purification of Avd mutants

Proteins were produced in soluble form in the *E. coli* strain BL21-AI using N-terminal OmpA bacterial secretion signal from *Bordetella avium* [26]. Proteins were purified by Ni-NTA affinity chromatography that yielded ~2 mg/L pure protein. According to gel-filtration analysis, in solution, both sbAvd-1 (51 kDa) and sbAvd-2 (53 kDa) showed tetrameric state (Additional file 3) similar to that of wtAvd (Avd expressed in *E. coli* 53 kDa, chicken Avd 60 kDa [26]). The slight decrease in molecular weight of Avd expressed in bacteria can be explained by the lack of glycosylation.

Determination of ligand-binding specificity of Avd forms by microplate assay

The binding specificity of proteins was analyzed by microplate assay, in which a set of different small molecules were used as a target molecules (Figure 4A). WtAvd was used as a negative control, and we detected no affinity towards the ligands except biotin. Importantly, sbAvd-1 and sbAvd-2 did not bind the proteins used as carriers for small molecules (bovine serum albumin (BSA) and human serum albumin (HSA)). However, these modified Avds showed clear binding to testosterone and progesterone.

The binding to the surface-immobilized ligands was inhibited by pre-incubating the proteins with 10 µM D-biotin (Figure 4B). Free biotin significantly inhibited the binding of sbAvd-1 to steroids, indicating notable affinity towards biotin. However, in the case of sbAvd-2 the binding of

testosterone and progesterone was not affected by biotin, showing a clear decrease in biotin-binding affinity.

Biosensor analyses of steroid-binding Avds

The kinetic constants of testosterone- and biotin-binding to steroid-binding proteins were determined with surface plasmon resonance (SPR) analysis. The purified sbAvd-1 and sbAvd-2 bound to a testosterone-BSA-coated sensor chip with similar affinities (Table 1), whereas wtAvd showed no binding to the testosterone surface. Furthermore, testosterone-binding was inhibited with varying testosterone concentrations (0.75-50 μM, data not shown). In the case of sbAvd-1, the 50% inhibition was achieved between 5-10 μM testosterone, which is consistent with the determined testosterone surface binding affinity. Interestingly, the 50% inhibition was already achieved in 750 nM testosterone concentration in the sbAvd-2 measurements. This result suggests a much higher binding affinity towards free testosterone than that measured towards surface-immobilized testosterone.

The biotin binding of the steroid-binding proteins was analyzed by the biotin-coated sensor chip. The comparison of sbAvd-1 and sbAvd-2 revealed that the biotin-binding affinity decreased almost 500-fold due to the modification of the loop between β-strands 3 and 4 (Table 1). This result appears to be consistent with the previous study in which mutation of T35A alone decreased the biotin-binding affinity of wtAvd approximately 200-fold [20].

The specificity of the steroid-binding was analyzed by binding competition using the following steroids: testosterone, dehydroepiandrosterone sulfate (DHEAS), androstenedione, estradiol, and dihydrotestosterone (DHT). In addition, the binding to the surface was also competed for by biotin. SbAvd-1 was found to be cross-reactive with androgens highly similar to testosterone (DHEAS and

androstenedione), and had noticeable affinity towards biotin (Figure 5A). Interestingly, DHT showed less efficient inhibition compared to testosterone, suggesting lower binding affinity towards this steroid form. The binding of sbAvd-2 to testosterone was found to be most efficiently inhibited by testosterone and DHEAS (Figure 5B), whereas the inhibition caused by biotin was clearly weaker than that in the case of sbAvd-1. This result is consistent with the microplate analysis results and with the binding kinetic constants determined with SPR analysis.

Interaction analysis of steroid-binding Avds by MRFS

Molecular recognition force spectroscopy (MRFS) [27] was used to study the interaction between sbAvds and testosterone on a single molecule level. An atomic force microscopy (AFM) tip was functionalized with a single testosterone molecule [28] and repeatedly approached and retracted from the sbAvd-1- or sbAvd-2-coated surface (Figure 6A). The binding forces were measured in force-distance cycles, whereby the deflection (force) of the cantilever was recorded as a function of the tip-sample distance. For evaluation, 1,000 force-distance cycles were recorded, and probability density functions (pdf) were generated from the detected interaction forces [29]. The tip-tethered testosterone was found to form a complex with the sbAvds; the retraction led to a downward bending of the cantilever until a particular force was reached resulting in the rupture of the bond between testosterone and sbAvd (Figure 6B and 6C). The most probable unbinding force [29] was found to be similar for both of the testosterone-sbAvd complexes: 40 pN at a constant pulling velocity of 600 nm/s. From 1,000 recorded force-distance cycles, sbAvd-1 showed 183 detected interactions and sbAvd-2 showed 215 events.

To prove the specificity of the measured interactions, control experiments were performed. For these experiments, free testosterone was injected into the measuring solution to preoccupy the binding sites of the steroid-binding protein immobilized on the surface. The retraction curve

identical to the approaching curve was detected (insets of Figure 6B and 6C), which indicated the total inhibition of the binding by free testosterone. Additionally, when the binding sites were preoccupied with free testosterone, the number of binding events dropped down to 67 from 183 detected interactions in the case of sbAvd-1. In case of sbAvd-2 the number of interactions dropped from of 215 to 21 after addition of free testosterone.

Differential scanning calorimetry reveals the high stability of the engineered proteins

The high thermal stability of wtAvd (temperature transition midpoint (T_m) = 85.5 °C) was preserved in the steroid-binding mutants; these proteins showed similar T_m values compared to wtAvd (sbAvd-1 T_m value of 80.6 °C and sbAvd-2 T_m value of 82.5 °C) (Table 1, Additional file 4). Ligand binding often stabilizes a protein and raises the T_m value. As expected, the addition of biotin stabilized wtAvd effectively, increasing the T_m value almost 40 °C. This T_m indicates a very high binding affinity.

The effect of biotin-binding on the protein stability of steroid-binding proteins was much smaller than in the case of wtAvd, thus showing decreased biotin affinity. In fact, sbAvd-2 showed a negligible increase in T_m in the presence of biotin (ΔT_m = 0.5 °C). The presence of testosterone only slightly increased the T_m values of sbAvds (ΔT_m = 0.6-0.9 °C). Actually, a similarly small increase was detected in the ΔT_m value of wtAvd when testosterone was added. However, because the testosterone-binding of wtAvd was not detected in other analyses (SPR and microplate analysis) the mechanism of stabilization remains unclear.

Discussion

In the present study, Avd proteins were displayed on the M13 phage in a monovalent form fused with the C-terminal region (aa residues 198-406) of the coat protein pIII. The crucial requirement

for the functional display on phage is successful expression of the protein in *E. coli*. As has been earlier reported Avd protein can be efficiently expressed in a soluble form in *E. coli* [26]. The close relative of Avd, streptavidin, has been displayed on bacterial phages as a pVIII fusion [30, 31]. However, to our best knowledge, chicken Avd has not been previously displayed on phages for screening purposes.

To evaluate the different modes for Avd display, we generated constructs expressing pIII fused to Avd and constructs that expressed free Avd subunits in addition to the pIII fusions. Based on the expression analysis, binding properties and selection experiments, the construct that expressed both fusion and free Avd showed an enhanced assembly of functional, tetrameric Avd on the phage compared to the Avd-pIII fusion alone (Additional file 2). However, the construct that expressed pIII-fused Avd was chosen for use in the selection system for the generation of gene libraries because the use of only one Avd gene in the protein display would ensure that the selected proteins could be assembled to homotetramers.

The potential of the developed platform for screening novel Avd-based receptors was investigated. Five residues in the loop between β-strands 1 and 2 of Avd were randomized to generate a population of diverse Avd genes. The gene population was screened for novel binding properties using the phage display method [23]. We observed the enrichment of Avd phages binding to testosterone-coupled BSA. The Avd variant sbAvd-1, which carried the sequence N12, R13, M14, N15, H16, was further analyzed and found to bind free testosterone with an affinity similar to BSA-conjugated testosterone. Moreover, we detected that the biotin-binding affinity of sbAvd-1 was not completely diminished; it was still high enough to inhibit sbAvd-1 from binding to testosterone-BSA (Table 1, Figures 4 and 5). This finding confirmed that testosterone occupies the same binding

site in Avd as does biotin. Based on the cross-creativity measurements, sbAvd-1 also binds other steroid hormones, such as progesterone and DHEAS (Figures 4 and 5).

SbAvd-1 was further engineered to decrease biotin binding and to improve steroid specificity. From the phage display selections, the sbAvd-1-derived protein, named sbAvd-2, with a mutated loop between β-strands 3 and 4 was captured. This mutant showed similar or slightly decreased binding affinity towards immobilized testosterone, whereas the biotin-binding affinity was clearly decreased (Table 1, Figures 4 and 5). We also noticed that sbAvd-2 bound more tightly to free testosterone than did sbAvd-1 (Figure 5). Significantly, sbAvd-2 preferred steroids as a ligand over biotin; this finding is important when considering the applications for biofluids, in which biotin is often present in relatively high concentrations.

This study and our preliminary results from experiments with a number of different target ligands (Hiltunen S, Riihimäki TA *et al.*, unpublished data) suggest that Avd-based receptors for various different small molecules can be tailored. These Avd-based receptors may be valuable tools for diagnostic use in the future.

Conclusions

The current study provides a promising platform for the selection of tailored ligand-binders evolved from the Avd scaffold in the monovalent pIII protein display. The Avd scaffold has characteristics that are beneficial in protein engineering, such as high thermal and chemical stability, simple folding and an optimal structure for small ligands. Furthermore, Avd can be modified rather freely without major change in the fold [18]. Novel Avds that can simultaneously bind multiple ligands could become next-generation molecular tools for clinical and diagnostic applications [22, 32]. Importantly, novel Avd-based receptors could be used in applications that require harsh conditions.

Methods

Construction of phagemid vectors

All of the basic recombinant DNA methods were performed essentially as previously described [33]. Appropriate restriction sites were added to the cDNA of the Avd core sequence [34] by PCR using the primers Avd_NheI_5' and Avd_NotI_3' (Additional file 5). PCR products were first subcloned into the pCR®2.1-TOPO plasmid by TOPO TA-cloning (Invitrogen) and the plasmids were transformed into *E. coli* TOP10 cells. Plasmids were isolated from colonies that contained the inserts based on the blue-white screening. Avd fragments were cut out from the pCR®2.1-TOPO-plasmid using the NheI and NotI restriction enzymes and ligated into the phagemid vector (pBluescript SK+ derived phagemid, Research Center of Finland, Biotechnology, Espoo, Finland). The cDNA of the Avd was cloned into the phagemid vector as an N-terminal fusion to the C-terminal domain (amino acids 198-406) of the minor phage coat protein III. In the Avd/Avd-pIII constructs, the coding sequence of the free Avd was subcloned using the primers Avd_NheI_5' and Avd_AscI_stop_3' (Additional file 5). The Avd-pIII expression cassette was generated by subcloning a fragment amplified from the primers Avd_SfiI_5' and Avd_NotI_3' (Additional file 5). The nucleotide sequences of the Avd constructs were verified by sequencing on an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems) according to the protocols recommended by the manufacturer (ABI PRISM BigDye Terminator Cycle Sequencing Kit v.1.1, Applied Biosystems).

Amplification of phage particles

All of the basic phage display methods were performed essentially as previously described [35]. Phagemid vectors with the Avd insert were transformed into chemically competent *E. coli* XL1-Blue cells (Stratagene, La Jolla, CA) with the heat shock method. Phage stocks of the different Avd display constructs were made from individual colonies picked from the transformation plates into

super broth (SB) medium supplemented with the appropriate antibiotics and glucose. The bacterial cultures were infected with the helper phage (10^{12} pfu/ml) VCS-M13 (Stratagene, LaJolla, CA) for amplification of Avd phages. Phages were PEG precipitated and analyzed by SDS-PAGE and western blotting following immunostaining with the polyclonal rabbit α-avd IgG (University of Oulu, 1:5000) and the monoclonal mouse anti-pIII IgG (Biosite, Sweden, 1:2000) antibodies.

Functionality test of Avd-displaying phages

The functionality of the Avd phages was tested by panning the phages on surfaces coated with BSA conjugated to HABA (HABA-BSA) [36]. As a positive control for selection, the Avd mutant N118M was also displayed on the phages. Panning was performed essentially as previously described [35]. For elution, vigorous shaking in 100 mM hydrochloric acid containing 10 µM D-biotin (Biochemica, Fluka, 14400) was performed. In total, three selection rounds were performed. After each round the integrity of the Avd expression units was analyzed by restriction enzyme digestions of the phagemid DNA samples. Importantly, both single and double Avd and Avd(N118M) construct phages were competed against each other. Equivalent amounts of phages displaying Avd or Avd(N118M) mutant were mixed and biopanned. The ratio between constructs was determined with DNA sequencing.

Construction of Avd L1,2 library

To construct an Avd DNA library, amino acids N12, D13, L14, G15, and S16 in the loop between beta strands 1 and 2 (L1,2) were randomized. The libraries were constructed essentially as previously described [35]. A nucleic acid fragment of 105 base pairs was amplified with the primers Avd_NheI_5' and Loop 1-2 _R1_3' (Additional file 5) using wt Avd cDNA as template. Parallel to this fragment, a nucleic acid fragment with 357 base pairs was PCR-amplified with the primers Loop 1-2_R2_5' and Avd_NotI_3' (Additional file 5) using wt Avd as a template. The PCR

strategy is presented in Additional file 6. The desired amplification products were separated by agarose gel electrophoresis and isolated from the gel using the Nucleo Spin Extract II kit (Macherey-Nagel) according to the manufacturer's instructions. PCR products were combined in the second amplification step in the presence of the PCR primers Avd_NheI_5' and Avd_NotI_3' (Additional file 5). This amplification resulted in a DNA fragment of 462 base pairs. The fragment was isolated from the gel and cut with the restriction enzymes *NheI* and *NotI* (Fermentas) followed by purification using the Nucleo Spin Extract II kit (Macherey-Nagel). Fragments were ligated into the phagemid vector and the resulting ligation product was transformed into electrocompetent cells of the *E. coli* strain XL1-Blue (Stratagene) by electroporation. Transfected bacteria were infected with the VCS-M13 helper phage (Stratagene) and phages were harvested from the culture and purified with PEG precipitation. The amount of phage particles was determined by titration.

Selecting steroid-binders from the Avd L1,2 library

The Avd L1,2 library was panned against the steroid hormone testosterone. As a negative control, wtAvd-displaying phages were also panned against testosterone. NUNC immunosorp plates were coated with a testosterone-BSA conjugate (Sigma, T-3392, 1 µg). Phages were preincubated in the BSA-coated wells to prevent non-specific binding. Three to four selection cycles were conducted, and the stringency of washing conditions was increased every panning round to decrease nonspecific binding. In the first panning round wells were washed three times with phosphate buffered saline containing 0.05% Tween (PBS-Tween) and five times with phosphate buffered saline (PBS). Phages were eluted by vigorous shaking for 10 minutes in 100 mM hydrochloric acid containing 10 µM D-biotin (Biochemica, Fluka, 14400). Biotin was used for elution because it was probable that after randomization of the loop L1,2 the resulting Avd would still have a rather high affinity towards biotin. In the final panning round 10 µM testosterone (Steraloids Inc., USA) was used in addition to 100 mM hydrochloric acid for elution. The eluted phage solutions were

neutralized with 2 M Tris. Eluted phages were used to infect *E. coli* XL1-Blue cells and aliquots of the infected bacteria were plated to quantify the amount of eluted phages. Phages were amplified and purified as described earlier. Precipitated phages were then used for the next round of selection. After every panning round, results were verified by sequencing (20 sequences), and the number of phage particles was determined by phage titration.

Generation and screening of the SbAvd-1 L3,4 library

An SbAvd-1-derived DNA library was constructed and ligated into the phagemid as described above. In the library amino acids T35, A36, V37, and T38 in the loop between beta strands 3 and 4 (L3,4) were randomized; thus, in the construction of the library the primers Avd_NheI_5' and 3_4R_1_3' and primers 3_4R_2_5' and Avd_NotI_3' (Additional file 5) were used in the PCR. SbAvd-1 cDNA was used as a template. Ligated phagemid was transformed into electrocompetent cells of *E. coli* strain XL1-Blue (Stratagene) by electroporation. Transfected bacteria were infected with the VCS-M13 helper phage (Stratagene) and phages were harvested from the culture and purified with PEG precipitation. The amount of phage particles was determined by titration.

The sbAvd-1 L3,4 library was biopanned against the steroid hormone testosterone similarly as described above, with some exceptions. NUNC immunosorp plates were coated with a testosterone-BSA conjugate (Sigma, T-3392, 200 ng). Phages were preincubated in BSA-coated wells before the panning procedure. Four selection cycles were performed. Phages were eluted by vigorous shaking for 10 minutes in 100 mM hydrochloric acid containing 35 µM testosterone (Steraloid Inc., USA). After each panning round, results were verified by sequencing (20 sequences), and the number of phage particles was determined by phage titration. Additionally, every panning round was screened for binders by anti-M13 as previously described [37]. For the ELISA, (NUNC) immunosorp plates were coated with a testosterone-BSA conjugate (Sigma, T-3392, 1 µg), wells were blocked with 5%

milk solution, and detection was performed with Anti-M13/HRP (GE Healthcare) and read with a microplate reader (Bio-Rad 680 XR).

Production and purification of Avd mutants

For biochemical analyses, the proteins were produced in *E. coli* strain BL21-AI (Invitrogen) using expression vector pET101/D (Invitrogen) [26]. After sonication (Sonics & Materials Vibra Cell™) and DNaseI (New England Bio Labs) treatment of *E. coli* cells, the purification of the proteins was conducted using Ni-NTA affinity chromatography according to the instructions of the manufacturer (QIAGEN).

The oligomeric state of the proteins was assayed with fast protein liquid chromatography (FPLC) gel-filtration using an ÄKTApurifier™ HPLC equipped with a Superdex 200 10/300 GL column (Tricorn, Amersham Biosciences, GE Healthcare). The column was calibrated using the gel-filtration mixture (thyroglobulin, γ-globulin, ovalbumin, myoglobin, and vitamin B12; Bio-Rad Laboratories) as a molecular-mass standard. Sodium phosphate buffer (20 mM, pH 7.4) with 1M NaCl and, 20mM imidazole was used as the liquid phase. Protein samples of 90–193 µg in a volume of 500 µl were used in the analysis.

DSC measurements of wtAvd and steroid-binding Avds.

Proteins (0.225 mg/ml) were analyzed in sodium phosphate buffer (20 mM, pH 7.4), containing 20 mM imidazole and 1 M NaCl. D-biotin (Biochemica, Fluka, 14400) and testosterone (Steraloids Inc., USA) were diluted with the measurement buffer to a final concentration of 50 µM. All solutions were degassed prior to measurements to avoid air bubbles. An automated capillary VP-DSC instrument (GE Healtcare, MicroCal, Northampton, USA) was used to measure the stability of the proteins with or without ligands. During the measurement, protein samples were heated from 20

°C to 130 °C at a scanning rate of 120 °C/h. Feedback mode was set to low and the filler period was 8 s. Temperature transition midpoints (T_m) were recorded from the highest peaks and the calorimetric heat changes (ΔH) were calculated using the MicroCal Origin 7 software (GE Healtcare, MicroCal, Northampton, USA).

Determination of ligand-binding specificity of Avd forms by microplate assay

MaxiSorp F96 microplate wells (NUNC) were coated with 500 ng of conjugated ligand (HSA-conjugated progesterone, hydrocortisone, testosterone, and cholic acid (Technical Research Center of Finland, Espoo, Finland); or BSA-conjugated testosterone (A6958-000, Steraloids Inc., USA), and biotin (Jenni Leppiniemi, University of Tampere, Finland)), and with 500 ng of the carrier proteins HSA and BSA (A7906, Sigma) in 100 µl of PBS for 2 hours 37 °C. Carrier proteins were used as negative controls. Plates were washed three times with PBS-Tween, blocked with 0.5 % BSA-PBS for 30 min, and then washed again. Proteins (0.9 µg/ml) in sodium phosphate buffer (20 mM, pH 7.4) with 1M NaCl and 20mM imidazole, were added to the wells and incubated for 1h. A portion of wtAvd (Belovo, Bastogne, Belgium), sbAvd-1 and sbAvd-2 was preincubated with 10 µM D-biotin (Biochemica, Fluka, 14400). Bound Avd was detected by rabbit α-avd IgG (University of Oulu) and with alkaline phosphatase conjugated goat anti-rabbit IgG (A3937, Sigma). After adding the phosphatase substrate solution (1 mg/ml pNPP (S0942, Sigma) in 1 M diethanolamine pH 9.8, with 0.5 mM $MgCl_2$), the plates were read after 15 minutes at A405 with a microplate reader (Bio-Rad 680 XR).

Biosensor analyses of steroid-binding Avds

A BIAcore X optical biosensor (Biacore, Uppsala, Sweden) was used for the analysis of binding kinetics. Testosterone-BSA was coupled to the carboxymethylated dextran layer of a sensor chip using standard amine coupling chemistry (1000 RU, 40 µl/min flow rate). Samples of sbAvd-1, and

sbAvd-2 were diluted in 50 mM sodium phosphate containing 1 M NaCl, and the same buffer was used in the measurements. The binding of the sbAvd-1 and sbAvd-2 samples on testosterone-BSA coated chips was measured and the kinetic constants were determined from the measurements performed with different protein concentrations using the BiaEvaluation software according to the manufacturer's instructions. WtAvd was used as a negative control.

The binding of steroid-binding proteins to the testosterone-BSA surface was competed with free steroid hormone molecules (testosterone, DHEAS, androstenedione, estradiol, and DHT (Steraloids Inc., USA)) and free biotin (Biochemica, Fluka, 14400) to evaluate the specificity of binding. Additionally, steroid-binding was more closely detected by measuring the binding in the presence of varying concentrations (0.75 μM-50 μM) of inhibiting testosterone.

For the determination of biotin-binding kinetics, a sensor chip was prepared as follows: diaminoethylene was first attached to the surface using a mixture containing 0.2 M 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC) and 0.05 M N-hydroxysuccinimide (NHS) in water. Second, to introduce amino groups to the surface, 1 M ethylenediamine (Fluka 03550) in water was applied. Finally, 5 mM biotin N-succinimidyl ester (Biochemica, Fluka, 14405) in 50% DMSO was injected on the surface (40 μl/min flow rate, ~130 RU; please note that the determination of the bound mass is not very accurate in case of small molecules because the immobilization can change the physicochemical properties of the surface).

Molecular Recognition Force Spectroscopy experiments

Molecular recognition force spectroscopy (MRFS) was used to study the interaction of the produced proteins with testosterone and biotin. Steroid-binding proteins were covalently bound to modified mica sheets via lysines as previously described [38]. Testosterone was coupled to the AFM tip via

the heterobifunctional Fmoc-PEG-NHS crosslinker as previously described [28], resulting in a covalent attachment of testosterone via a flexible spacer (Figure 6A).

All MRFS experiments were performed on a Pico SPM I (Agilent Technologies, Santa Clara, CA). All modified cantilevers used had nominal spring constants between 10–30 pN/nm (Veeco Instruments, Santa Barbara, CA). The effective spring constants of the cantilevers were determined by the thermal noise method [39, 40]. Force-distance cycles were completed using a z-range of 200 and 300 nm. Sweep durations were adjusted between 0.25 and 4 s. During one data set of 1000 force-distance curves, the lateral tip position was changed (a few hundred nm) about every 100 curves to ensure that the binding events were statistically reasonable. The specificity of the binding was proved by adding free testosterone (200 nM) into the measuring solution and incubating for approximately 1 h to block the ligand-binding sites of the proteins.

Abbreviations: AFM, atomic force microscopy; BSA, bovine serum albumin: BTN, D-biotin; DHEAS, dehydroepiandrosterone sulphate; DHT, dihydrotestosterone; DMSO, dimethyl sulfoxide; EDC, 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride; HABA, 4'-hydroxyazobenzene-2-carboxylic acid; HSA, human serum albumin; MRFS, molecular recognition force spectroscopy; NHS, N-hydroxysuccinimide; PBS, phosphate buffered saline; PEG, polyethylene glycol; sbAvd, steroid-binding avidin; SB, super broth; SDS-PAGE, sodium-dodecyl sulfate polyacrylamide gel electrophoresis; wtAvd, wild type avidin

Author contributions

TAR conceived the study, constructed the Avd libraries, performed most of the experiments, and contributed to the writing of the manuscript; SH participated in the construction of the Avd libraries and in the experimental work and contributed to the writing of the manuscript; JAEM participated

in the experimental work and contributed to the writing of the manuscript; MR, AE and PH performed MRFM measurements and contributed to the writing of the manuscript; HRN, MSK and KT conceived the study and supervised the work; VPH conceived and designed the study, supervised the work and contributed to the writing of the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We thank the laboratory staff of VTT (Technical Research Center of Finland) Biotechnology, Espoo, Finland, for excellent technical assistance. We thank MSc. Antti Tullila for kindly providing us HSA-conjugates. We also thank the staff at the Molecular Biology group, Department of Biological and Environmental Science, University of Jyväskylä for their assistance. We thank MSc. Sampo Kukkurainen for help in molecular modeling. We would also like to thank MSc. Kaisa Helttunen, Professor Kari Rissanen and Dr. Juhani Huuskonen for their help in the preparation of the HABA-BSA conjugate. We thank Hong Chang for participation in the experimental work. The study was financially supported by the Academy of Finland (115976, 121236); the National Technology Agency of Finland (BioFace 40055/08); European Micro and Nano Technology support program (FFG 421695); Pirkanmaa Hospital District and Tampere Graduate Program in Biomedicine and Biotechnology.

References

1. Hoogenboom HR: **Selecting and screening recombinant antibody libraries.** Nat Biotechnol 2005, **23**:1105-1116.

2. Hoess RH: **Protein design and phage display.** Chem Rev 2001, **101**(10):3205-3218.

3. Sarikaya M, Tamerler C, Jen AK, Schulten K, Baneyx F: **Molecular biomimetics: nanotechnology through biology.** Nat Mater 2003, **2**:577-585.

4. Binz HK, Amstutz P, Pluckthun A: **Engineering novel binding proteins from nonimmunoglobulin domains.** Nat Biotechnol 2005, **23**:1257-1268.

5. Skerra A: **Alternative non-antibody scaffolds for molecular recognition.** Curr Opin Biotechnol 2007, **18**:295-304.

6. Werner RG: **Economic aspects of commercial manufacture of biopharmaceuticals.** J Biotechnol 2004, **113**:171-182.

7. Hudson PJ, Souriau C: **Engineered antibodies.** Nat Med 2003, **9**:129-134.

8. Saerens D, Ghassabeh GH, Muyldermans S: **Single-domain antibodies as building blocks for novel therapeutics.** Curr Opin Pharmacol 2008, **8**:600-608.

9. Nygren PA, Skerra A: **Binding proteins from alternative scaffolds.** J Immunol Methods 2004, **290**:3-28.

10. Gebauer M, Skerra A: **Engineered protein scaffolds as next-generation antibody therapeutics.** Curr Opin Chem Biol 2009, **13**:245-255.

11. Beste G, Schmidt FS, Stibora T, Skerra A: **Small antibody-like proteins with prescribed ligand specificities derived from the lipocalin fold.** Proc Natl Acad Sci U S A 1999, **96**:1898-1903.

12. Schlehuber S, Skerra A: **Tuning ligand affinity, specificity, and folding stability of an engineered lipocalin variant -- a so-called 'anticalin' -- using a molecular random approach.** Biophys Chem 2002, **96**:213-228.

13. Laitinen OH, Nordlund HR, Hytönen VP, Kulomaa MS: **Brave new (strept)avidins in biotechnology.** Trends Biotechnol 2007, **25**:269-277.

14. Green NM: **Avidin.** Adv Protein Chem 1975, **295**:85-133.

15. Repo S, Paldanius TA, Hytönen VP, Nyholm TK, Halling KK, Huuskonen J, Pentikäinen OT, Rissanen K, Slotte JP, Airenne TT, Salminen TA, Kulomaa MS, Johnson MS: **Binding Properties of HABA-Type Azo Derivatives to Avidin and Avidin-Related Protein 4.** Chem Biol 2006, **13**:1029-1039.

16. Livnah O, Bayer EA, Wilchek M, Sussman JL: **Three-dimensional structures of avidin and the avidin-biotin complex.** Proc Natl Acad Sci U S A 1993, **90297**:5076-5080.

17. Rosano C, Arosio P, Bolognesi M: **The X-ray three-dimensional structure of avidin.** Biomol Eng 1999, **16**:5-12.

18. Laitinen OH, Hytönen VP, Nordlund HR, Kulomaa MS: **Genetically engineered avidins and streptavidins.** Cell Mol Life Sci 2006, **63**:2992-3017.

19. Nordlund HR, Laitinen OH, Hytönen VP, Uotila ST, Porkka E, Kulomaa MS: **Construction of a dual chain pseudotetrameric chicken avidin by combining two circularly permuted avidins.** J Biol Chem 2004, **279**:36715-36719.

20. Hytönen VP, Nordlund HR, Hörhä J, Nyholm TK, Hyre DE, Kulomaa T, Porkka EJ, Marttila AT, Stayton PS, Laitinen OH, Kulomaa MS: **Dual-affinity avidin molecules.** Proteins 2005, **61**:597-607.

21. Riihimäki TA, Kukkurainen S, Varjonen S, Hörhä J, Nyholm TKM, Kulomaa MS, Hytönen VP: **Construction of chimeric dual-chain avidin by tandem fusion of the related avidins**. PLoS One 2011, **6**:e20535.

22. Nordlund HR, Hytönen VP, Hörhä J, Maatta JA, White DJ, Halling K, Porkka E, Slotte JP, Laitinen OH, Kulomaa MS: **Tetravalent single chain avidin: From subunits to protein domains via circularly permuted avidins.** Biochem J 2005, **392**:485-491.

23. Smith GP: **Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface.** Science 1985, **228**:1315-1317.

24. Hoogenboom HR, Griffiths AD, Johnson KS, Chiswell DJ, Hudson P, Winter G: **Multi-subunit proteins on the surface of filamentous phage: methodologies for displaying antibody (Fab) heavy and light chains.** Nucleic Acids Res 1991, **19**:4133-4137.

25. Määttä JA, Airenne TT, Nordlund HR, Jänis J, Paldanius TA, Vainiotalo P, Johnson MS, Kulomaa MS, Hytönen VP: **Rational modification of ligand-binding preference of avidin by circular permutation and mutagenesis.** Chembiochem 2008, **9**:1124-1135.

26. Hytönen VP, Laitinen OH, Airenne TT, Kidron H, Meltola NJ, Porkka E, Hörhä J, Paldanius T, Määttä JA, Nordlund HR, Johnson MS, Salminen TA, Airenne KJ, Ylä-Herttuala S, Kulomaa MS: **Efficient production of active chicken avidin using a bacterial signal peptide in Escherichia coli.** Biochem J 2004, **384**:385-390.

27. Hinterdorfer P, Dufrene YF: **Detection and localization of single molecular recognition events using atomic force microscopy.** Nat Methods 2006, **3**:347-355.

28. Wildling L, Hinterdorfer P, Kusche-Vihrog K, Treffner Y, Oberleithner H: **Aldosterone receptor sites on plasma membrane of human vascular endothelium detected by a mechanical nanosensor.** Pflugers Arch 2009, **458**:223-230.

29. Baumgartner W, Hinterdorfer P, Schindler H: **Data analysis of interaction forces measured with the atomic force microscope.** Ultramicroscopy 2000, **82**:85-95.

30. Avrantinis SK, Stafford RL, Tian X, Weiss GA: **Dissecting the streptavidin-biotin interaction by phage-displayed shotgun scanning.** Chembiochem 2002, **3**:1229-1234.

31. Sidhu SS, Weiss GA, Wells JA: **High copy display of large proteins on phage for functional selections.** J Mol Biol 2000, **296**:487-495.

32. Leppiniemi J, Määttä JA, Hammaren H, Soikkeli M, Laitaoja M, Jänis J, Kulomaa MS, Hytönen VP: **Bifunctional avidin with covalently modifiable ligand binding site.** PLoS One 2011, **6**:e16576.

33. Sambrook J, Fritsch EF, Maniatis T: **Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY,** 1990.

34. Gope ML, Keinänen RA, Kristo PA, Conneely OM, Beattie WG, Zarucki-Schulz T, O'Malley BW, Kulomaa MS: **Molecular cloning of the chicken avidin cDNA.** Nucleic Acids Res 1987, **15**:3595-3606.

35. Barbas CF, III, Burton DR, Scott JK, Silverman GJ: **Phage Display, A Laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY,** 2001.

36. Hofstetter H, Morpurgo M, Hofstetter O, Bayer EA, Wilchek M: **A labeling, detection, and purification system based on 4-hydroxyazobenzene-2-carboxylic acid: an extension of the avidin-biotin system.** Anal Biochem 2000, **284**:354-366.

37. Kingsbury GA, Junghans RP: **Screening of phage display immunoglobulin libraries by anti-M13 ELISA and whole phage PCR.** Nucleic Acids Res 1995, **23**:2563-2564.

38. Kamruzzahan AS, Kienberger F, Stroh CM, Berg J, Huss R, Ebner A, Zhu R, Rankl C, Gruber HJ, Hinterdorfer P: **Imaging morphological details and pathological differences of red blood cells using tapping-mode AFM.** Biol Chem 2004, **385**:955-960.

39. Hutter JL, Bechhoefer J: **Calibration of atomic-force microscope tips.** Review of Scientific Instruments 1993, **64**:1868-1873.

40. Butt HJ, Jaschke M: **Calculation of thermal noise in atomic force microscopy.** Nanotechnology 1995, **6**:1-7.

41. Livnah O, Bayer EA, Wilchek M, Sussman JL: **Three-dimensional structures of avidin and the avidin-biotin complex.** Proc Natl Acad Sci USA 1993, **90**:5076-5080.

42. Humphrey W, Dalke A, Schulten K: **VMD: visual molecular dynamics.** J Mol Graph 1996, **14**:33-8, 27-8.

Figure legends

Figure 1. Schematic presentation of the Avd display expression constructs. (A) The phagemid constructs for Avd and Avd(N118M) display, in which Avd protein is produced solely as a fusion with pIII. (B) The phagemid constructs for Avd/Avd-pIII and Avd(N118M)/Avd(N118M)-pIII display, in which the *pelB* signal sequence is used for secretion of the Avd-pIII fusions and the free Avd. The cloning sites used are shown by vertical arrows.

Figure 2. Immunoblot analysis of Avd phages with anti-Avd and anti-pIII antibodies. (A) The location of the Avd-pIII fusion protein recognized by anti-Avd is indicated by the upper arrowhead. The theoretical mass of the Avd-pIII fusion protein is 38 kDa. The produced free Avd (~14 kDa) is indicated with the lower arrowhead. (B) From the immunoblot analyzed with anti-pIII antibody the full-length pIII expressed from the helper phage (VCSM13) can be seen migrating at ~60 kDa (the upper arrowhead). The Avd-pIII fusion is indicated by the middle arrowhead. In addition, some proteolytically truncated Avd-pIII forms were detected (~33 kDa; lower arrowhead). Lane 1: Avd-pIII phage; lane 2: Avd(N118M)-pIII phage; lane 3: Avd/Avd-pIII phage; lane 4: Avd(N118M)/Avd(N118M)-pIII phage. The molecular weights of standard proteins are shown as kilodaltons on the left side of each blot.

Figure 3. Three-dimensional structure of wtAvd with the loops chosen for random mutagenesis marked. One subunit of tetrameric wtAvd with bound biotin (PDB: **2AVI**) [16] is shown in the figure. The randomized amino acid residues are N12, D13, L14, G15, and S16 (shown in blue) in the loop between β-strands 1 and 2, and T35, A36, V37, and T38 (shown in purple) in the loop between β-strands 3 and 4.

Figure 4. Determination of ligand-binding specificity of sbAvd-1 and sbAvd-2 proteins by microplate analysis. The binding of sbAvd-1 and sbAvd-2 to a set of different small ligands was detected using polyclonal anti-avidin antibody as a probe (A) and the effect of free biotin (10 µM) to ligand-binding was analyzed (B).

Abbreviations used in the figure: PBS, Phosphate buffered saline; HSA, human serum albumin; PRO, HSA-conjugated progesterone; HYD, HSA-conjugated hydrocortisone; TES, HSA-conjugated testosterone; CHO, HSA-conjugated cholic acid; TES-B, BSA-conjugated testosterone; BTN, BSA-conjugated D-biotin; BSA, bovine serum albumin

Figure 5. Inhibition analysis of sbAvd-1 and sbAvd-2 proteins by the SPR method. The binding of the sbAvd-1 and the sbAvd-2 to a CM5 sensor chip functionalized with testosterone-BSA was measured in the presence of 50 μM inhibitors. (A) The binding of the sbAvd-1 protein was totally inhibited by dehydroepiandrosterone, androstenedione, and biotin. (B) The binding of the sbAvd-2 protein was totally inhibited by dehydroepiandrosterone or testosterone, but not biotin. This result is due to the markedly decreased affinity of the protein towards biotin.

Samples: Protein sample, green; protein with estradiol, dark blue; protein with DHT, olive; protein with testosterone, black; protein with DHEAS, blue; protein with androstenedione, brown; protein with biotin, red

Figure 6. Analysis of sbAvd protein-testosterone interaction by MRFS method (A) A schematic representation of the experimental assembly used in the analyses. Testosterone was tethered to the AFM tip using a flexible PEG crosslinker. The steroid-binders sbAvd-1 and sbAvd-2 were covalently bound to the mica surface via a short homobifunctional spacer. (B) The force-distance cycle of the sbAvd-1 – testosterone interaction showing an unbinding event. The typical non-linear shape of the event results from the elastic properties of the PEG linker. (C) The force-distance cycle showing a sbAvd-2 – testosterone bond dissociation. Insets in (B) and (C) represent force-distance cycles in which the protein-testosterone interaction is inhibited with free testosterone.

Tables

Table 1. Kinetic parameters and the determined thermostability of sbAvds
The kinetic parameters determined by SPR analysis. The biotin binding affinity of avidin is too high to be determined by SPR. The transition midpoint of thermal unfolding (T_m) and the calorimetric heat of unfolding were determined by DSC. Delta T_m represents the increase of T_m in the presence of 50 μM ligand.

Protein	Ligand	SPR			DSC		
		k_a (1/Ms)	k_d (1/s)	K_D (M)	T_m (˚C)	ΔT_m (˚C)	$\Delta H \times 10^4$ (cal/mol)
wtAvd	-	-	-	-	85.5	-	5.4
wtAvd	Btn	n.d.	n.d.	n.d.	123.2	37.7	12.8
wtAvd	Tes	-	-	no binding	86.2	0.7	5.6
sbAvd-1	-	-	-	-	80.6	-	5.2
sbAvd-1	Btn	4.2×10^5	5.6×10^{-4}	1.4×10^{-9}	83.2	2.6	6.4
sbAvd-1	Tes	1.0×10^3	9.5×10^{-3}	9.0×10^{-6}	81.5	0.9	5.7
sbAvd-2	-	-	-	-	82.5	-	4.6
sbAvd-2	Btn	1.0×10^3	6.8×10^{-4}	6.6×10^{-7}	83.0	0.5	4.2
sbAvd-2	Tes	813	8.5×10^{-3}	1.1×10^{-5}	83.1	0.6	4.8

Additional files

Additional file 1

Structural comparison of Avd-BTN and Avd-HABA complexes.

3D-structures of Avd complexed with biotin (A) and HABA (B). These two ligands have different hydrogen bonding interactions with the amino acids at the binding site. (A) Hydrogen bonds formed between the ureido ring of biotin and Avd are shown here as dashed green lines. (X-ray crystallographic structure (PDB **2AVI**)) [41]. (B) A significant difference between biotin and HABA binding to Avd is seen in the interaction of the ligand with asparagine 118. Although other Avd key residues interacting with the BTN ureido ring also interact with HABA, there is no hydrogen bonding partner for N118 in HABA. Again, hydrogen bonds formed between the carboxyl group of HABA and Avd are shown as dashed green lines. The loop between *β*-strands 3 and 4 was not resolved in the avd-HABA –complex (coordinates kindly provided by Prof. Oded Livnah). The figure was made with the VMD program [42].

Additional file 2

Sequencing results and input from the control selections.

The percentage of wt Avd and Avd(N118M) mutant sequences after sequencing analysis from the different rounds of HABA selection. The amount of input phages is shown as colony forming units (cfu) per milliliter of culture.

Additional file 3

Gel-filtration chromatograms of sbAvd-1 and sbAvd-2 proteins.

The chromatograms of steroid-binding proteins determined at wavelength of 280 nm by gel-filtration. Besides the main peak, there is also a small peak of oligomeric form of the protein in the chromatogram of sbAvd-1 protein (gray curve). It is a typically observation also in the case of wtAvd [26]. The chromatogram of sbAvd-2 protein is shown with black curve. The tetrameric form of the protein is dominant in both samples.

Additional file 4

The effect of ligand-binding to the stability of wtAvd, sbAvd-1 and sbAvd-2 proteins.

DSC thermograms were obtained from the protein sample (0.225 mg/ml) by scanning temperature range of 20 °C to 130 °C with heating rate of 120 °C/min. The analysis was conducted in the absence and presence of ligands (50 μM).

Additional file 5

The cloning primers used in the study.

The restriction enzyme cleavage sites are indicated in *italics*.

Additional file 6

A schematic presentation of the construction strategy of the Avd L1,2 library.

For 1-2 loop library a nucleic acid fragment of 105 base pairs was amplified (Step 1, A) with the primers Avd_NheI_5' and Loop 1-2 _R1_3' (schematically referred in the figure as Avd_mutant_3') using wtAvd cDNA as a template. Parallel to this, a nucleic acid fragment with 357 base pairs, was PCR-amplified (Step 1, B) with the primers Loop 1-2_R2_5' (schematically referred in the figure as Avd_combine_5') and Avd_NotI_3', also using wtAvd as a template. Amplified fragments were combined in a second amplification step in the presence of PCR primers Avd_NheI_5' and Avd_NotI_3', wherein a DNA fragment of 462 base pairs was obtained.

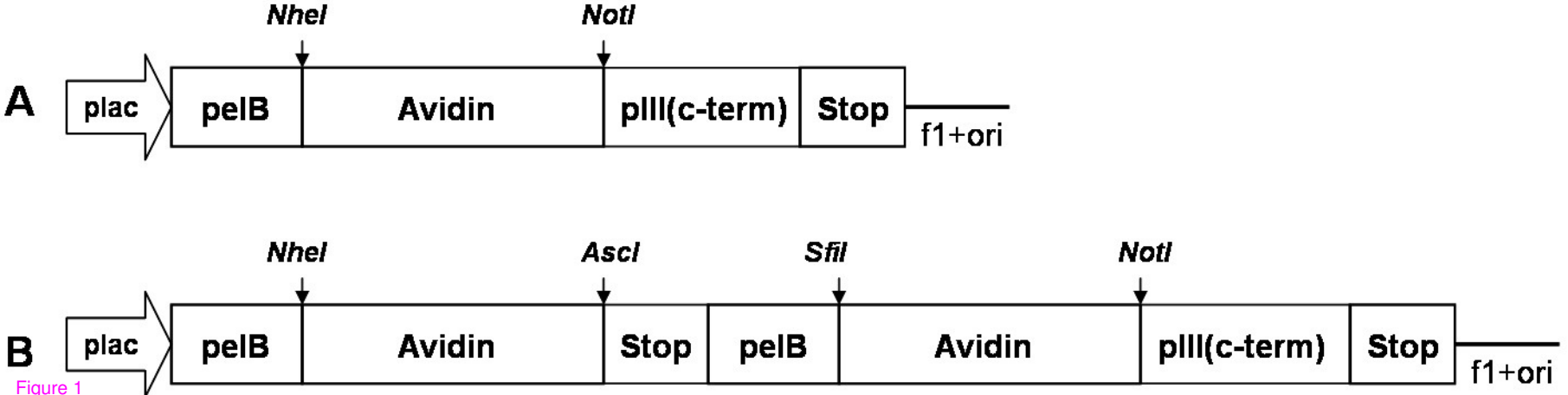

A
NheI
NotI
plac
pelB
Avidin
pIII(c-term)
Stop
f1+ori
B
NheI
AscI
SfiI
NotI
plac
pelB
Avidin
Stop
pelB
Avidin
pIII(c-term)
Stop
f1+ori

Figure 1

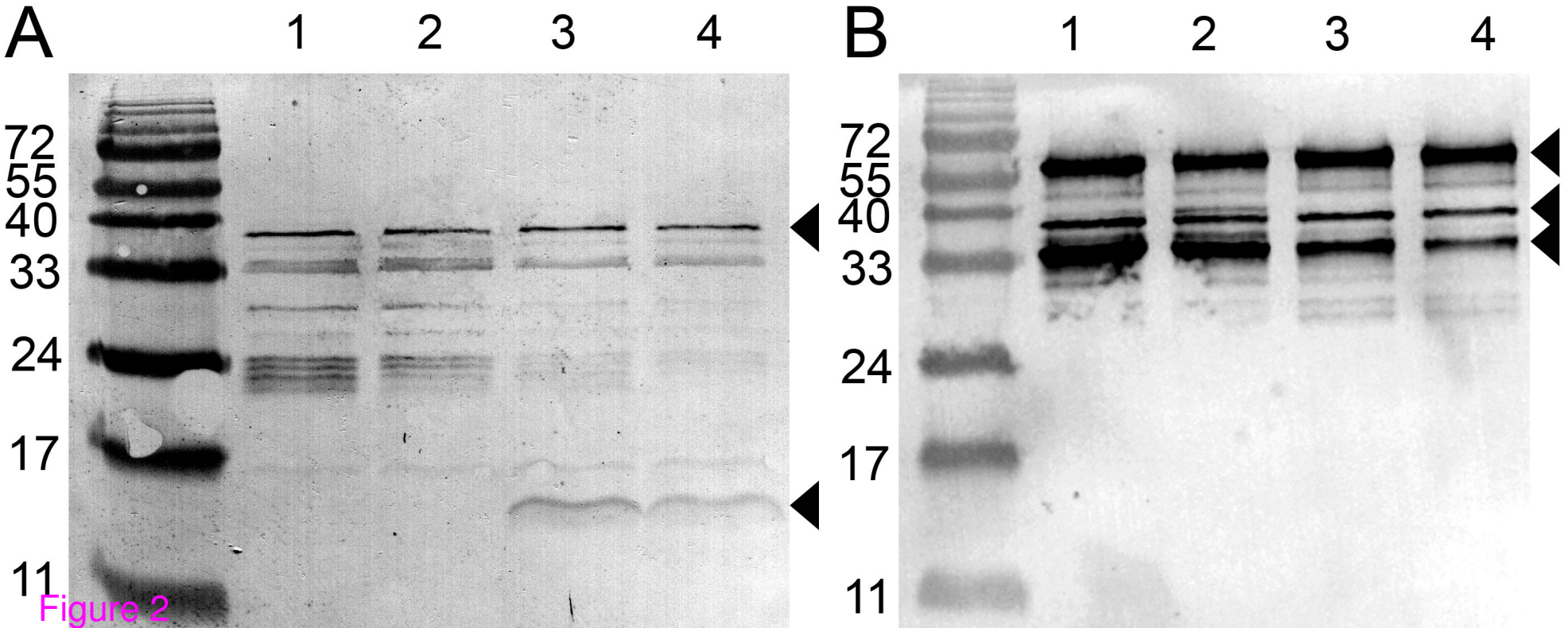


Figure 2

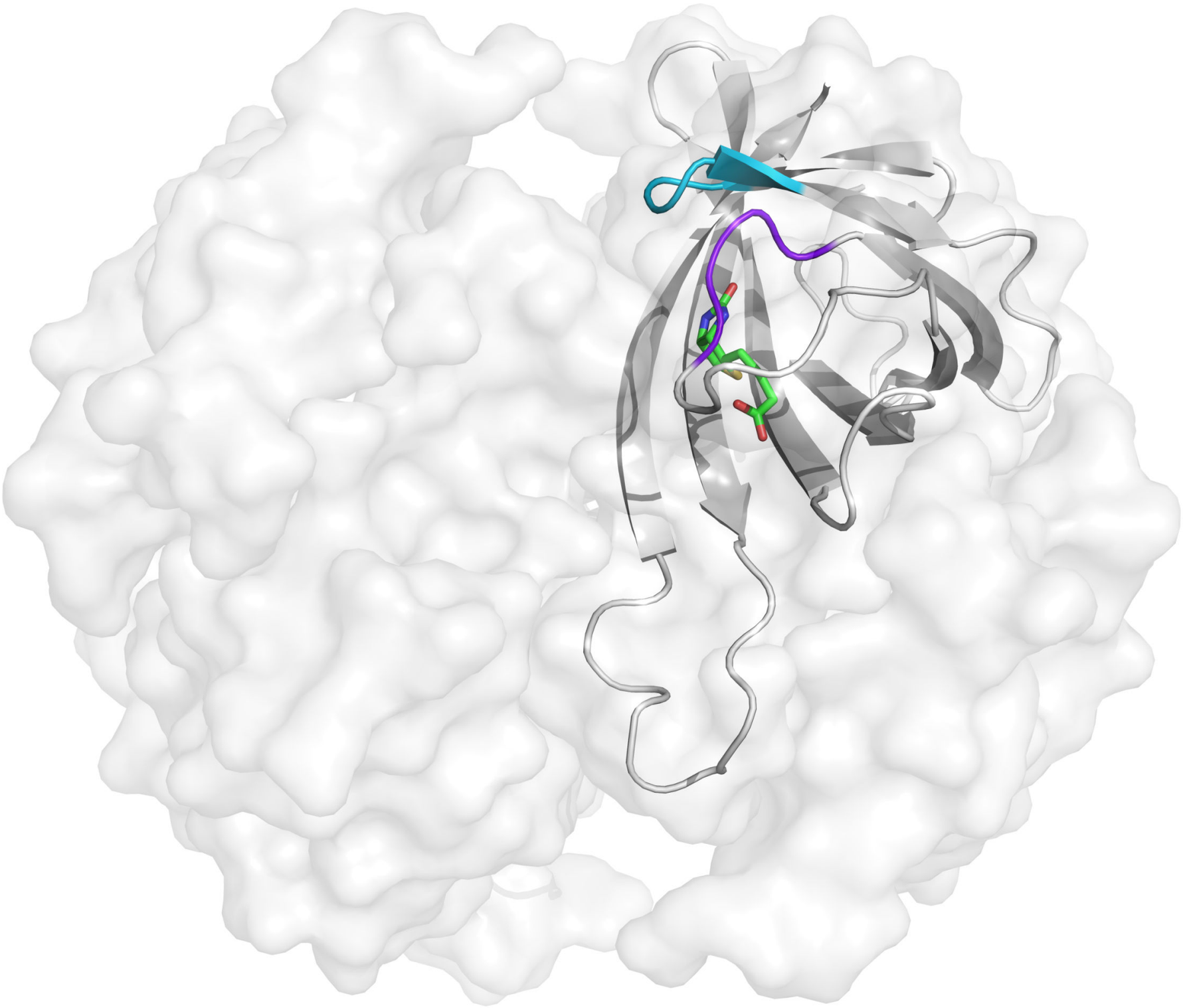
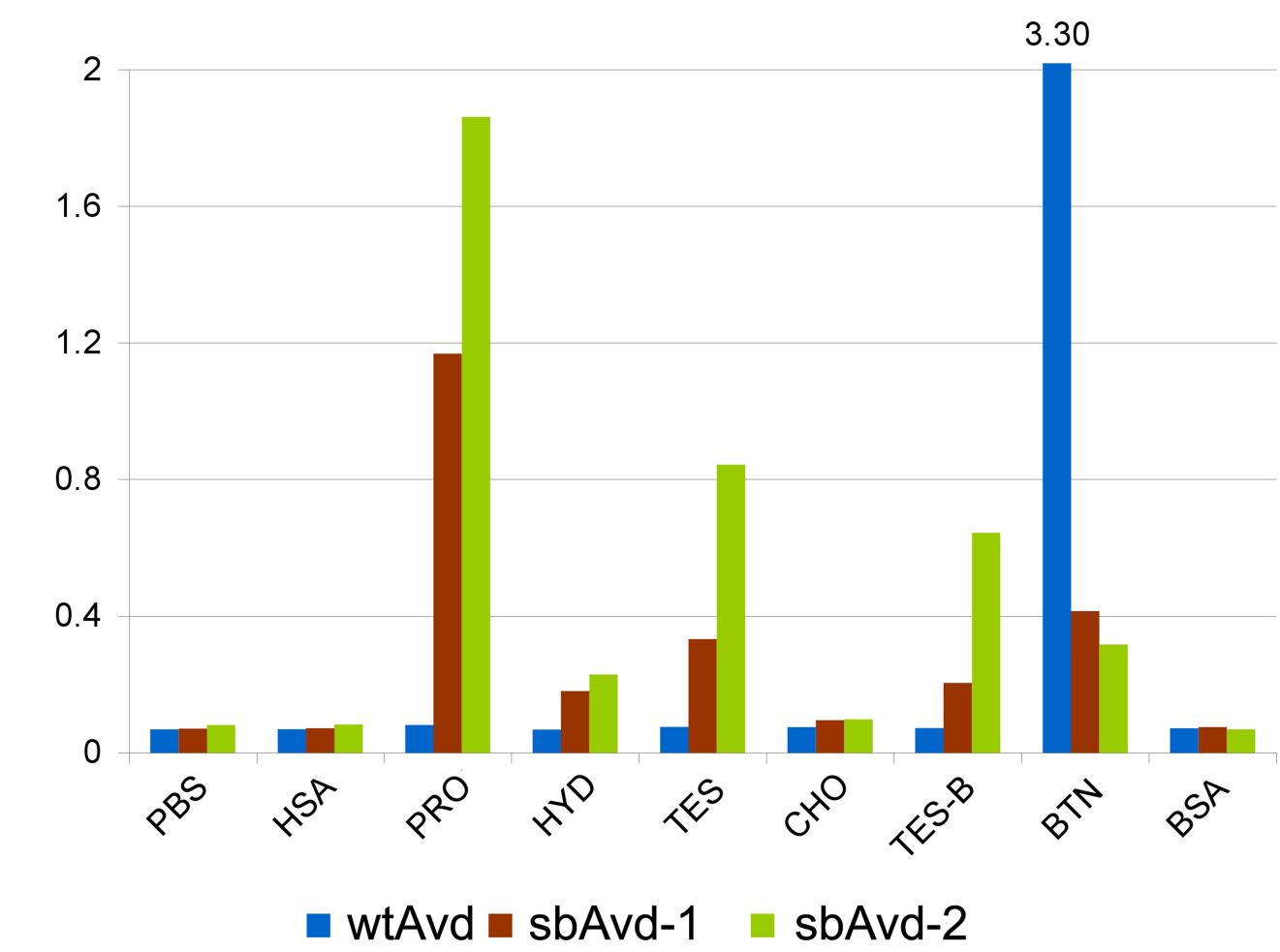


Figure 3


3.30
2
1.6
1.2
0.8
0.4
0
PBS
HSA
PRO
HYD
TES
CHO
TES-B
BTN
BSA
wtAvd
sbAvd-1
sbAvd-2

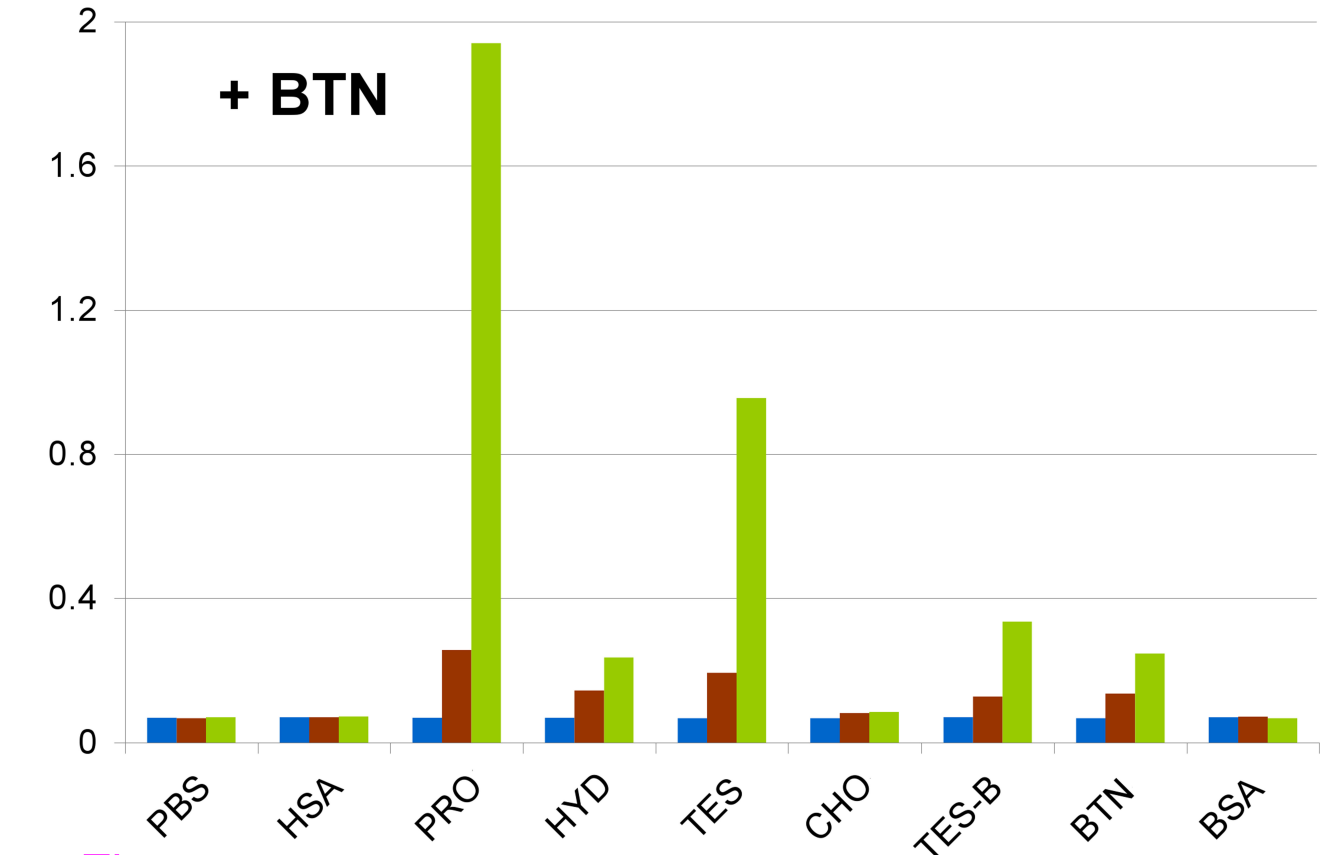
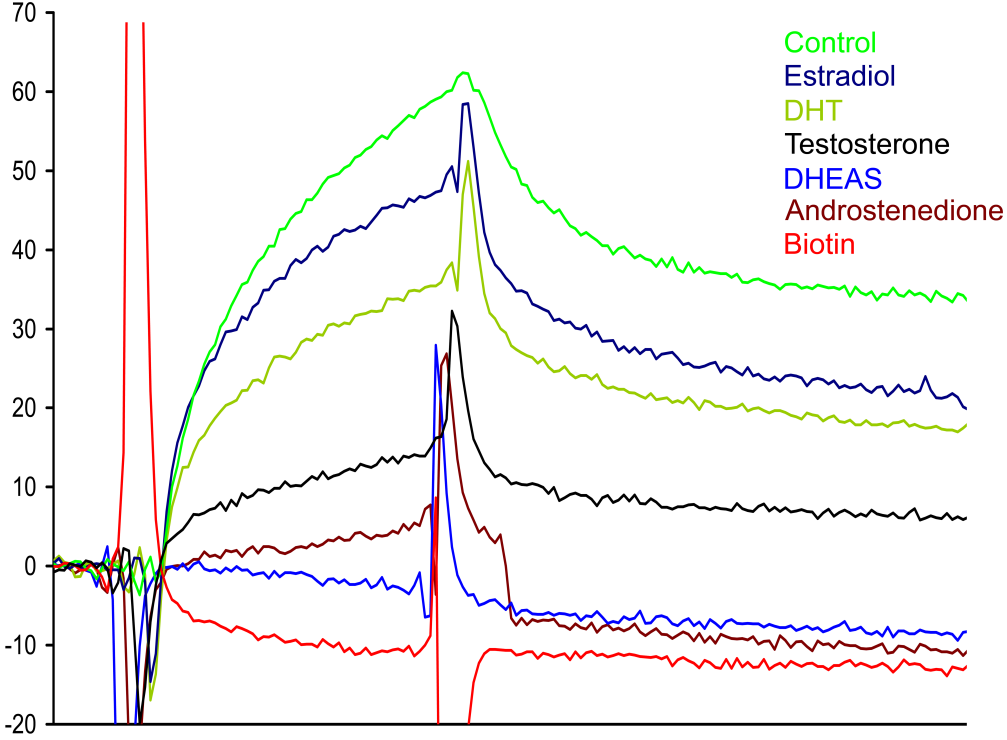
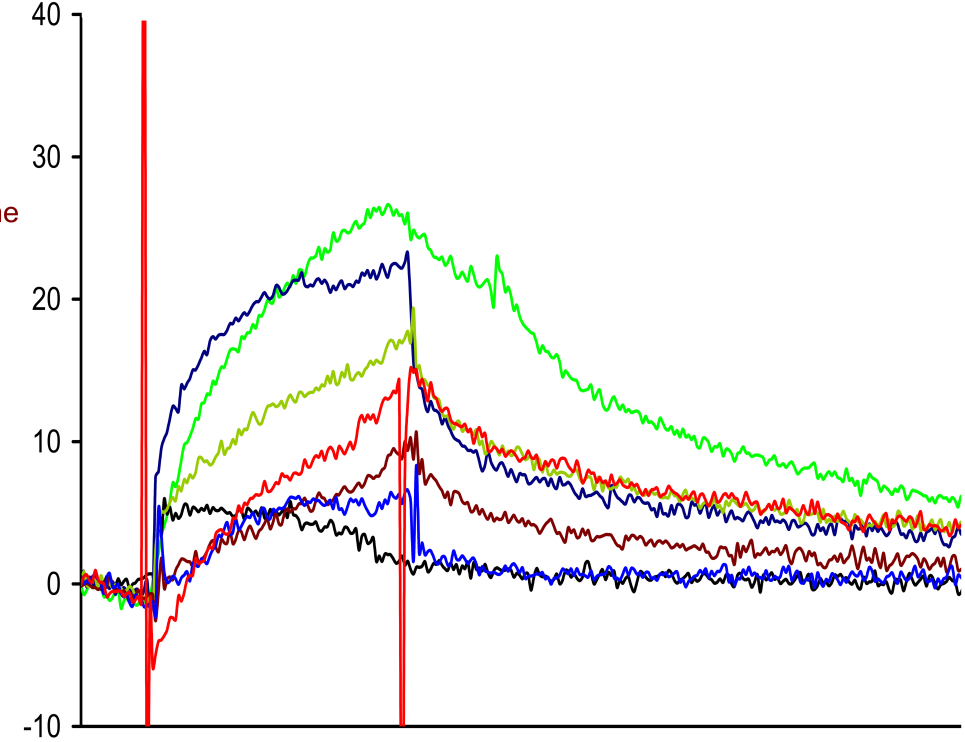

+ BTN
2
1.6
1.2
0.8
0.4
0
PBS
HSA
PRO
HYD
TES
CHO
TES-B
BTN
BSA

Figure 4


70
60
50
40
30
20
10
0
-10
-20
Control
Estradiol
DHT
Testosterone
DHEAS
Androstenedione
Biotin

A


40
30
20
10
0
-10

B

Figure 5

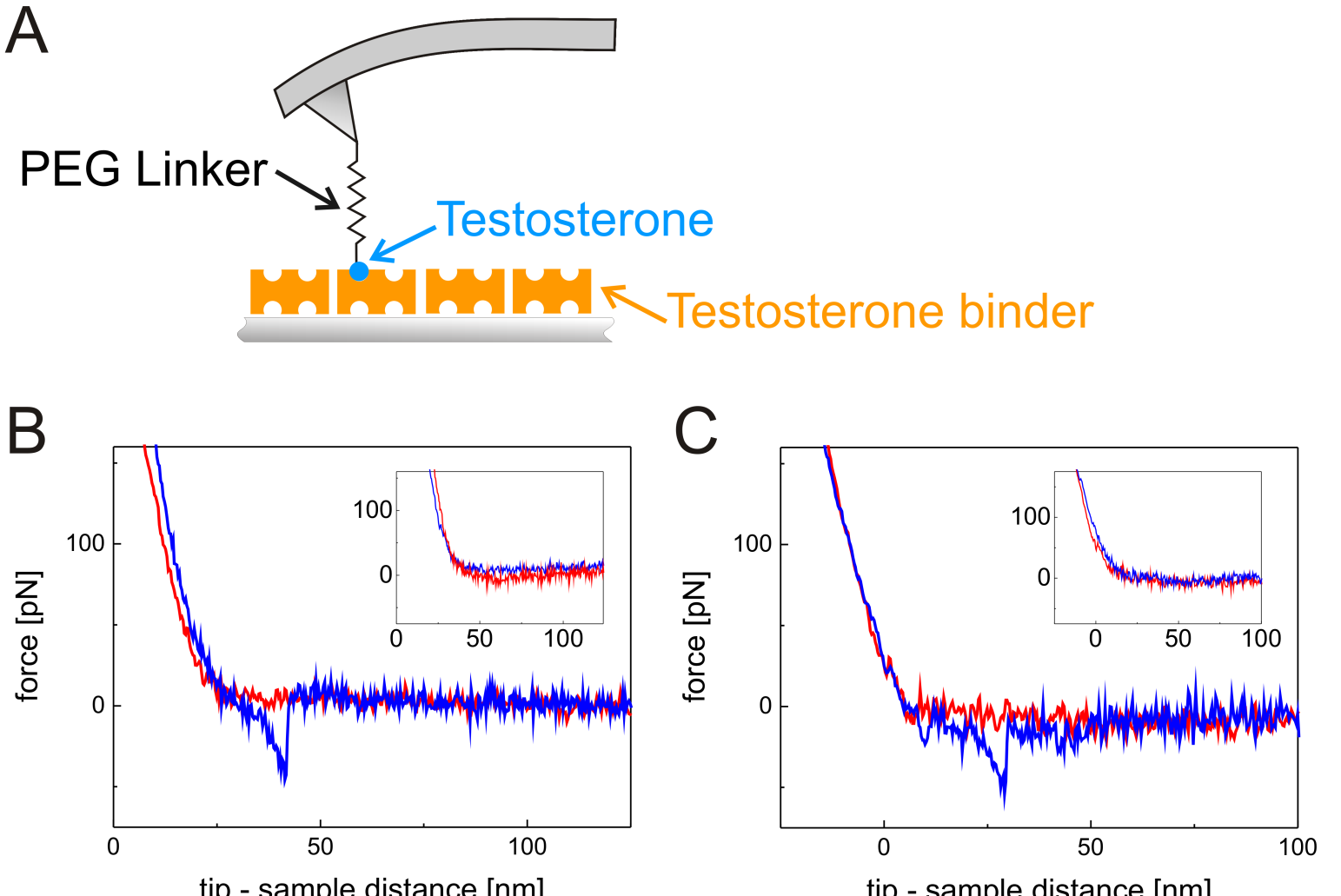

A
PEG Linker
Testosterone
Testosterone binder
B
force [pN]
100
0
0
50
100
tip - sample distance [nm]
100
0
0
50
100
C
force [pN]
100
0
0
50
100
tip - sample distance [nm]
100
0
0
50
100

Figure 6

Additional files provided with this submission:

Additional file 1: Additional file 1.png, 712K
http://www.biomedcentral.com/imedia/1926332432560144/supp1.png
Additional file 2: Additional file 2.doc, 26K
http://www.biomedcentral.com/imedia/2100180528560145/supp2.doc
Additional file 3: Additional file 3.tif, 879K
http://www.biomedcentral.com/imedia/2094349168560144/supp3.tiff
Additional file 4: Additional file 4.tif, 1856K
http://www.biomedcentral.com/imedia/9233958356014519/supp4.tiff
Additional file 5: Additional file 5.doc, 31K
http://www.biomedcentral.com/imedia/1303715701560145/supp5.doc
Additional file 6: Additional file 6.png, 122K
http://www.biomedcentral.com/imedia/6588496756700243/supp6.png